A Link to the Math

Connections Between Number Theory and Other Mathematical Topics

Martin Kreh

Between Number Theory and Other Mathematical Topics
the Math Connections A Link to

UV Universitätsverlag Hildesheim

Martin Kreh

A Link to the Math

Connections Between Number Theory and Other Mathematical Topics

Martin Kreh

A Link to the Math

Connections Between Number Theory and Other Mathematical Topics

Universitätsverlag Hildesheim

Hildesheim 2018

Impressum

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Universitätsverlag Hildesheim
Universitätsplatz 1
31141 Hildesheim

https://www.uni-hildesheim.de/bibliothek/publizieren/open-access-universitaetsverlag/

Erstausgabe Hildesheim 2018
Redaktion, Satz und Gestaltung: Martin Kreh

Der Beitrag ist abrufbar unter:
http://dx.doi.org/10.18442/805

Zitierempfehlung:

Kreh, Martin (2018). *A Link to the Math. Connections Between Number Theory and Other Mathematical Topics.* Hildesheim: Universitätsverlag Hildesheim. DOI: http://dx.doi.org/10.18442/805 (Open Access)

Martin Kreh

A Link to the Math

Connections Between Number Theory and Other Mathematical Topics

Dissertation submitted to the University of Hildesheim
in partial fulfillment of the requirements for the degree of
Dr. rer. nat. in mathematics

February 13, 2017
revised September 2, 2017

The
main title
is based on the
Zelda adventure

A
Link
to the Past

that
is now
available since
$12^{h(-12)}{}_{\omega(12)}{}^{\tau(12)}$ days

Acknowledgements

There are countless (but not uncountably many) people without whose this thesis would not be as it is.

I would like to thank all teachers, either at school or at university, who taught and helped me in my mathematical education. In particular I would like to thank Dr. Hanns-Joachim Strunck, Klaus Günthert and Hans-Joachim Hoinka, who influenced me a lot during my time at school and who laid the foundations for my studies. I thank all lecturers at the different universities of whom I attended lectures. I especially thank Prof. Dr. Knut Smoczyk and Prof. Dr. Stefan Wewers, whose very different styles of lecture had taught me much in my first years at university.

We are on a learning process our whole life. During my time at the University of Hildesheim I have learned a lot through research, teaching and discussions with colleagues. I would like to thank my friends and colleagues at the Institute of Mathematics and Applied Computer Science for the nice working atmosphere and all the things I learned. In particular I thank Dr. Jan-Hendrik de Wiljes for useful discussions and corrections as well as for his support in our work routine. It has always been a pleasure to work with him.

I would especially like to thank my advisor Prof. Dr. Jürgen Sander. He showed me some nice problems to attack and also gave me the freedom to work on the problems that came to my mind. His kind of supervision and his assistance were very helpful for both my teaching and my research.

There are many other mathematicians whose work or help guided me through my studies and research. I would like to thank all of them, in particular Dr. Marcos Soriano, who helped me with a problem in Chapter II.3, Prof. Dr. Carsten Elsner, who provided the proof of a lemma in Chapter II.4, and Dr. Ioulia Baoulina and Prof. Dr. Jörn Steuding, the co-authors of one of the articles whose content

is a part of Chapter II.1. I also like to thank the community of MathOverflow. I learned a lot both through answers of my own and foreign questions.

Additionally I would like to thank Prof. Dr. Karl Dilcher for being available as referee and Dr. Mario Müller for his assistance regarding the final publication of this thesis.

I would never be able to do mathematics all day long. For the past years it has always been nice to have friends and family who took my mind off work. I also thank my parents and my parents-in-law for their support during my studies and my writing time.

Especially I would like to thank my wife, Anja. She helped me whenever I needed diversion and mental support, and she bore my bad mood whenever a problem had put me down. Thank you for your encouragement and the great time that we had (and most hopefully will have) together.

Abstract

Number theory is one of the oldest mathematical areas. This is perhaps one of the reasons why there are many connections between number theory and other areas inside mathematics. This thesis is devoted to some of those connections.

In the first part of this thesis I describe known connections between number theory and twelve other areas, namely analysis, sequences, applied mathematics (i.e., probability theory and numerical mathematics), topology, graph theory, linear algebra, geometry, algebra, differential geometry, complex analysis, physics and computer science, and algebraic geometry. We will see that the concepts will not only connect number theory with these areas but also yield connections among themselves.

In the second part I present some new results in four topics connecting number theory with computer science, graph theory, algebra, and linear algebra and analysis, respectively.

First I consider the minimal set of a set $M \subset \mathbb{N}$, i.e., the smallest subset $A \subset M$ such that every number $m \in M$ can be reduced to a number $a \in A$ by striking away some digits. I will determine the minimal set for some arithmetically interesting sets and develop an algorithm for congruence classes.

In the next topic I determine the neighbourhood of the neighourhood of vertices in some special graphs. This problem can be formulated with generators of subgroups in abelian groups and is a direct generalization of a corresponding result for cyclic groups.

In the third chapter I determine the number of solutions of some linear equations over factor rings of principal ideal domains R. In the case $R = \mathbb{Z}$ this can be used to bound sums appearing in the circle method.

Lastly I investigate the puzzle “Lights Out” as well as variants of it. Of special interest is the question of complete solvability, i.e., those cases in which all starting boards are solvable. I will use various number theoretical tools to give a criterion for complete solvability depending on the board size modulo 30 and show how this puzzle relates to algebraic number theory.

Zusammenfassung

Zahlentheorie ist eine der ältesten mathematischen Gebiete. Möglicherweise ist dies der Grund dafür, warum es so viele Verbindungen zwischen Zahlentheorie und anderen Gebieten innerhalb der Mathematik gibt. Diese Arbeit beschäftigt sich mit einigen dieser Verbindungen.

Im ersten Teil dieser Arbeit werden bekannte Verbindungen zwischen Zahlentheorie und zwölf weiteren Gebieten, nämlich Analysis, Folgen, angewandte Mathematik (d.h. Wahrscheinlichkeitstheorie und numerische Mathematik), Topologie, Graphentheorie, lineare Algebra, Geometrie, Algebra, Differentialgeometrie, Funktionentheorie, Physik und Informatik, sowie algebraischer Geometrie, betrachtet. Wir werden sehen, dass die hier betrachteten Konzepte nicht nur Zahlentheorie mit den obigen Gebieten verbinden, sondern auch Verbindungen zwischen den Gebieten selbst liefern.

Im zweiten Teil stelle ich in vier Kapiteln eigene Resultate vor, die Zahlentheorie mit Informatik, Graphentheorie, Algebra sowie linearer Algebra und Analysis verbinden.

Zuerst betrachte ich die Streichungsmenge einer Menge $M \subset \mathbb{N}$, d.h. die kleinste Teilmenge $A \subset M$ so dass jede Zahl $m \in M$ durch Streichen von Ziffern zu einer Zahl $a \in A$ reduziert werden kann. Ich bestimme die Streichungsmenge für einige zahlentheoretisch interessante Mengen und entwickle einen Algorithmus für Kongruenzklassen.

Im nächsten Kapitel bestimme ich die Nachbarschaft der Nachbarschaft von Knoten in speziellen Graphen. Dieses Problem kann mit Hilfe von Erzeugern von Untergruppen in abelschen Gruppen formuliert werden und ist eine direkte Verallgemeinerung eines entsprechenden Ergebnisses für zyklische Gruppen.

In dritten Kapitel bestimme ich die Lösungsanzahl von linearen Gleichungen über Faktorringen von Hauptidealringen R. Für den Fall $R = \mathbb{Z}$ kann dies benutzt werden um Summen, die bei der Kreismethode auftauchen, abzuschätzen.

Zuletzt untersuche ich das Puzzle "Lights Out" sowie Varianten davon. Insbesondere wird die Frage nach der vollständigen Lösbarkeit betrachtet, d.h. es geht um die Fälle, in denen jede Startkonfiguration lösbar ist. Unter Benutzung verschiedener zahlentheoretischer Hilfsmittel beweise ich ein Kriterium für die vollständige Lösbarkeit, das von der Boardgröße modulo 30 abhängt, und zeige, wie "Lights Out" mit algebraischer Zahlentheorie zuammenhängt.

Introduction

Part I - Connections

Part II - Original Work

Appendix

Introduction

Introduction

Number theory is (one of) the oldest areas in mathematics.

This (or a similar) statement can be found in nearly every introductory text in number theory. Although one could argue that most of the tools used in contemporary number theory are more recent, some classical questions arose centuries ago. Maybe this is one of the reasons why there are so many connections between number theory and other mathematical topics. This thesis is devoted to these connections.

In my time as Ph.D. student I wrote several articles, most of which are part of this thesis. All of these articles have one thing in common: Aside from number theory they are located in (or at least touch on) some other area, namely analysis, linear algebra, algebra, graph theory and theoretical computer science. This fact resulted in the idea of examining connections between number theory and other mathematical areas more closely.

This thesis is divided into two main parts. In the first part I show some connections between number theory and other mathematical areas. All definitions, theorems, and concepts in that part are already known. These chapters serve as a small presentation in a nutshell. The second part of the thesis presents my own research. This consists of four topics that have been (or will be) published in five articles.

For the first part of this thesis one has to decide which other mathematical areas should be considered. Of course this raises the question

What is a mathematical area?

This question cannot really be answered. The standard way of classifying mathematical work is the MSC2010 classification, which has a total of more than 5000 entries with 63 first level entries, cf. [MSC]. One could define these 63 first level entries as "areas". But when looking at these one finds, for example, different first level entries for "commutative algebra" (13) and "associative rings and algebras" (16), two topics that could be considered as parts of the area "algebra". This means that, in my opinion, the MSC2010 classification is not suitable for identifying areas, at least in the context of this thesis.

There is another reason why the MSC2010 classification, as well as other classifications, e.g., the one used for articles in the Arxiv, are not practical. We will show this by means of two figures. These two figures show different "maps of maths", i.e., the connections between mathematical areas. The first map, shown in Figure 0.0.1, shows the connection when considering papers uploaded on the Arxiv between 1992 and 2014.

Figure 0.0.1: Connections between mathematical areas according to the Arxiv. Source: [Loo]

On this map, two areas are connected if a paper has been uploaded whose Arxiv classification included these two areas. The more such papers exist, the thicker the line is. In Figure 0.0.2, the same has been done with the MSC2010 classification and papers appearing at Zentralblatt Math and Mathematical Reviews. For more information on these "maps" see [Loo].

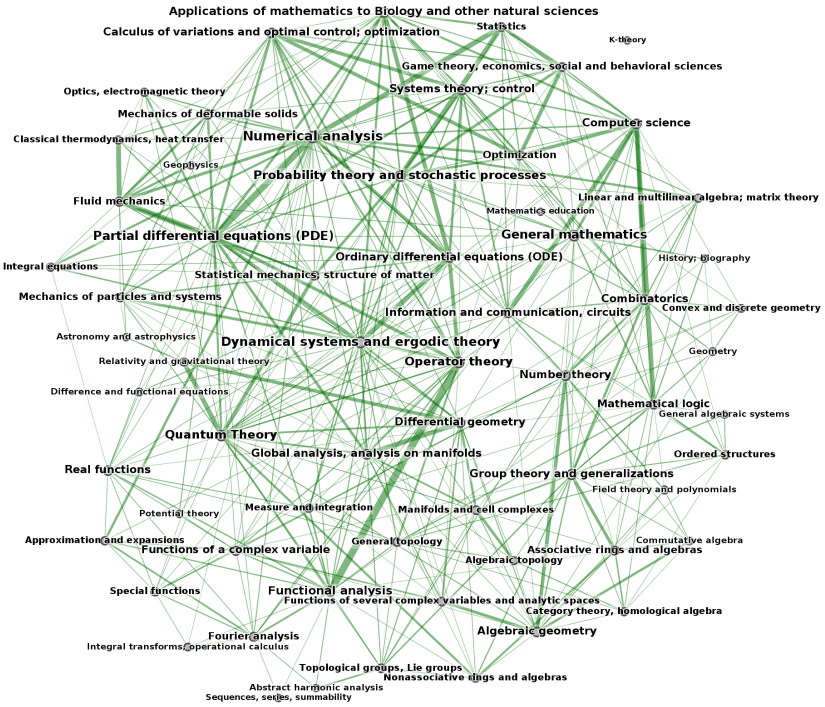

Figure 0.0.2: Connections between mathematical areas according to Zentralblatt Math and Mathematical Reviews. Source: [Loo]

These figures again show that these classifications allow too many areas. The other reason why this classification is not suitable is the fact that it does not show all existing connections. In Figure 0.0.1 there is, for example, no connection between "number theory" and "formal languages and automata theory". But in fact there is a really nice connection between these areas, as we will see in Chapter I.11. Thus I did not use the MSC2010 classification to identify the areas.

Another great presentation of mathematical fields (and mathematics itself) can be found in the Princeton Companion to Mathematics [GBGL08], which presents 26 mathematical branches. These branches are mostly too specialized to be suitable for our investigation. There are other introductory textbooks about the fields

of mathematics (for example [HH12], which includes too few areas for our purpose).

Thus, instead of using an existing apportionment of mathematics, I chose twelve areas which, in my opinion, cover a big part of mathematics and for which connections with number theory could be found (for example, it seems hard to find conections between number theory and category theory although there are papers like [FL05] and [Sch00]. We will also see a categorical construction in Chapter I.4, but we do not use methods from category theory, thus this is more a connection with topology than with category theory). The twelve areas that I will show connections with are

- (real) analysis
- sequences
- applied mathematics (i.e., probability theory and numerical mathematics)
- topology
- graph theory
- linear algebra
- geometry
- algebra
- differential geometry
- complex analysis
- physics and computer science
- algebraic geometry

A (slightly ironic) comment on why I chose exactly twelve other areas can be found in Appendix A.1.

In the twelve chapters of part 1, I will show one (or sometimes more) connection(s) of the respective topic with number theory. The goal is to mention concepts and theorems that either combine number theory with other areas or to examine theorems in number theory that can be proven with concepts or tools used in other areas. I will not present any proofs in that part (except in the case

when these proofs are illuminating to show some connections). In most cases I try to mention enough about the concepts so that they are understandable. There are, however, some exceptions. These exceptions result in the fact that I do not understand the concepts enough (or even not at all) to explain them. This is the case for étale cohomology, K-theory, arithmetic manifolds and all concepts in physics. I decided to include these topics here nonetheless, because I think that these belong in a thesis devoted to connections with number theory. In fact, most of the connecting concepts that I will present will appear in more than one chapter.

Although there are many research articles connecting number theory and other areas, there appears to be not much work in the treatment of such combinations, which justifies the first part of this thesis. There is one notable exception: The proceedings of the DIMACS workshop on "Unusual Applications of Number Theory" held in 2000, see [Nat04]. These proceedings contain some very nice connections between number theory and various other areas. We will see some of the concepts mentioned in these proceedings in the first part of this thesis.

In the second part of this thesis, I present my own research in four chapters.

In the first chapter I consider the problem of finding minimal sets for some subsets of the natural numbers. Here the minimal set of a set M of natural numbers is the smallest subset $A \subset M$ such that every number $m \in M$ can be reduced to a number $a \in A$ by deleting some of the digits of m. This problem was first introduced by Shallit in [Sha00a], but since then there has not been much progress. Here I consider some subsets of the natural numbers defined by arithmetic conditions (for example the set of natural numbers that can be written as a sum of two squares or values of some arithmetic functions). Further, I develop an algorithm that constructs the minimal set for congruence classes. In fact, this algorithm can be applied to a more general class of sets.

I show that minimal sets do not permit much structure, i.e., set-theoretic relations between two sets will, in general, not be passed on to the respective minimal sets. In addition to this, I show that measure-theoretic tools cannot help in determining the number of elements in minimal sets.

This topic can be viewed as a connection to theoretical computer science, since the definition of minimal sets originates from the theory of formal languages, as will become clearer later. These results have been published in the papers [Kre15b] and [BKS17], where the second is joint work with J. Steuding and I. Baoulina.

The next topic is about adding generators in abelian groups. In general, one does not expect much structure when adding generators. In 2013, J. W. Sander and T. Sander investigated the sumset of two atoms in cyclic groups A, i.e., for given $a, b, c \in A$ they explicitely computed the set of $u, v \in A$ such that $u + v = c$, where u is a generator of the subgroup $\langle a \rangle$ and v a generator of the subgroup $\langle b \rangle$, as well as the number of representations. I show that their result generalizes immediately to the case when A is an abelian group. Further I show that this result cannot hold for non-abelian groups in general.

This topic has connections to graph theory: The results can be applied to determine the neighbourhood of the neighbourhood of vertices in Cayley graphs. The results are published in [Kre15a].

In the third chapter I examine the problem of finding the number of solutions of linear congruences of the form $A\mathbf{x} \equiv \mathbf{0} \bmod m$ where A is an integer matrix and $m \in \mathbb{N}$. In fact this problem can be generalized, and I will determine the number of solutions of linear equations over factor rings of principal ideal domains with a certain finiteness condition by using the Smith normal form of the involved matrix. An explicit formula, depending only on the Smith normal form, as well as upper bounds are given.

This upper bound can be used in analytic number theory to bound the value of certain sums appearing in circle methods. Since the generalized problem is purely algebraic and we will use algebraic methods to determine the number of solutions, this topic is a connection between number theory and algebra.

The results are published in [Kre16].

Lastly, I investigate the puzzle "Lights Out" as well as some variants of it (in particular varying board size and number of colors). Here the player is given a quadratic board with illuminated buttons. The goal is to turn all lights out by pressing the correct combination of buttons. I discuss the complete solvability of such games, i.e., I am interested in the cases such that all starting boards can be solved. I will model the problem with basic linear algebra and develop a criterion for the unsolvability depending on the board size modulo 30. Further, I will discuss ways of handling the solvability that will rely on prime divisors of Lucas numbers, analysis and Diophantine approximation, and algebraic number theory, respectively. Hence this topic is a connection between linear algebra, analysis and number theory.

The results will be published in [Kre17].

At the end of each chapter of Part 2 I will present some open questions and discuss potential future work in the respective topic.

Before starting with the connections in Part 1 I will recall some basics. These will serve as basics for both parts of the thesis. Here I will restrict to definitions and theorems that we use later. I will further assume that the reader is acquainted with basic mathematics and elementary number theory.

In the appendix after Part 2 I will present some additional material. I will discuss interesting properties of the number 12 in Appendix A.1. In Appendix A.2 I will give an example of a concept that I will mention in Chapter I.10. This example will be shown because the only source in which I could find this example is not publicly accessible. In the appendix I will also show some computations and further examples for two articles from Part 2.

Most notation used in this thesis is standard. As usual in number theory, the symbol $\mathbb{N}$ denotes the natural numbers with $0 \notin \mathbb{N}$ and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. The set of primes is denoted by $\mathbb{P}$. The set $\mathbb{Z}_p$ denotes the p-adic integers, while $\mathbb{Z}/p\mathbb{Z}$ denotes the integers modulo p (and more generally, $\mathbb{Z}/m\mathbb{Z}$ denotes the integers modulo m).

To avoid confusion we will denote the greatest common divisor of two integers m, n by $\gcd(m, n)$. In this thesis, the divisor function is denoted by $\tau(n)$. We will denote the natural logarithm of a number x by $\log(x)$. Whenever the argument of log is a complex number, this denotes the principal value of the complex logarithm, i.e., the branch with imaginary part in $(-\pi, \pi]$. If z is a complex number, $\Re(z)$ and $\Im(z)$ denote the real and imaginary part, respectively.

If M is a set, we denote its complement (in a given natural superset of M, mostly $\mathbb{R}$ or $\mathbb{N}$) by M^c and its cardinality with $|M|$. When we need to distinguish between elements of M and M^n for some n (for example if M is a field, we want to distinguish between scalars and vectors), elements in M are denoted by italic letters, e.g., $x \in M$, while elements in M^n are denoted by boldface letters, e.g., $\mathbf{x} \in M^n$.

A complete list of symbols can be found in the appendix. I have tried to avoid using the same symbol for different concepts, but this was not always possible. In each case, it will be clear from the context which meaning the symbols have. Due to the structure of this thesis, the list of symbols is rather large. Thus the list of symbols is (not only) sorted alphabetically. The sorting is explained at the beginning of the list of symbols.

All results in Part 2 of this thesis are my own results except for those where I explicitely mention another author. Apart from the two figures in the introduction, all other graphics are done by myself, mostly using Ti*k*Z.

0

Basics

In the following sections we will state some basic facts and definitions that we will need in the rest of the thesis. The emphasis here is in understanding the concepts. We will not give any proofs of the results but instead give references where they can be found. Several paragraphs will end with a combined example of the most important definitions mentioned for the respective topic.

0.1 Topology and Geometry

We will start with some foundations in topology and geometry, in particular in differential geometry. We will just need some basic understanding of the concepts and will omit exact definitions if they require too much notation.

0.1.1 Topology

Definition 0.1.1 A **topological space** (M, τ) is a set M equipped with a topology τ, i.e., a set of subsets $\tau \subset \mathscr{P}(M)$ satisfying

- $\emptyset, M \in \tau$,
- if $S_1, S_2 \in \tau$, then $S_1 \cap S_2 \in \tau$,
- if $S_i \in \tau$ for $i \in \mathbb{N}$, then $\bigcup_{i \in \mathbb{N}} S_i \in \tau$.

Sets in τ are called **open**. A subset $K \subset M$ is called **closed** if $M \backslash K$ is open.

The **(topological) closure** of a set A in a topological space (M, τ) is the smallest closed subset $K \subset M$ that contains A. The closure of A is denoted by $\overline{A}$. The set A is called **dense** in M if $\overline{A} = M$. For example, $\mathbb{Q}$ is dense in $\mathbb{R}$ in the usual topology.

A topological space (M, τ) is called **connected** if there are no nonempty open sets $U_1, U_2 \in \tau$ such that $U_1 \cap U_2 = \emptyset$ and $U_1 \cup U_2 = M$.

Let (M_1, τ_1) and (M_2, τ_2) be two topological spaces. Then a map $f : M_1 \to M_2$ is called **continuous** if for all $U \in \tau_2$ we have $f^{-1}(U) \in \tau_1$, i.e., the preimage of any open set is open. Bijective continuous maps f such that the inverse of f is again continuous are called **homeomorphism**.

In some cases algebraic structures can be equipped with a topological structure that is compatible with the operations.

A **topological group** is a group G such that composition and inversion are both continuous (as maps from $G \times G$ to G or from G to G, respectively). A **topological ring** is a ring R such that addition and multiplication are both continuous (as maps from $R \times R$ to R).

For a ring R, the units R^* of R form a group. If R is a topological ring, R^* does not need to be a topological group when equipped with the topology of R. However, if we embed R^* in $R \times R$ via $x \mapsto (x, x^{-1})$ and consider R^* with the topology of $R \times R$ (i.e., the open sets of R^* are the sets $U \cap R^*$ where $U \subset R \times R$ is open), then this becomes a topological group. We call this topology the **IC topology** (**i**nversion **c**ontinuous) on R.

A topological space (M, τ) is called **Haussdorff space** if for any two distinct points $p_1, p_2 \in M$ there are open sets $U_1, U_2 \subset M$ such that $p_1 \in U_1, p_2 \in U_2$ and $U_1 \cap U_2 = \emptyset$.

If (M, τ) is a topological space, an **open covering** is a collection of open sets $U_i \subset M, i \in \mathbb{N}$ such that $\bigcup_{i \in \mathbb{N}} U_i = M$. An open covering $\mathfrak{C} = \{U_i\}$ is called **locally finite** if for any $p \in M$ there is an open set $U \subset M$ with $p \in U$ and such that the set $\{i : U \cap U_i \neq \emptyset\}$ is finite, i.e., any p is only covered with finitely many open sets from $\mathfrak{C}$.

If for any open covering $\{U_i\}$ there is a locally finite **refinement**, i.e., a locally finite covering $\{U'_j\}$ such that for any $j \in \mathbb{N}$ there is an $i \in \mathbb{N}$ with $U'_j \subset U_i$, then (M, τ) is called **paracompact**.

Let $\{U_i\}$ be an open covering of a topological space (M, τ). A **subcover** of $\{U_i\}$ is a covering $\{U'_j\}$ such that for any $j \in \mathbb{N}$ there is an $i \in \mathbb{N}$ with $U'_j = U_i$. A set A in a topological space (M, τ) is called **compact** if any open covering of A

has a finite subcover. (M, τ) is called **locally compact** if for any $p \in M$ there is a compact set K that includes an open set $U \subset M$ with $p \in U$. For example, the real numbers $\mathbb{R}$ with the usual topology are locally compact, while the set of rational numbers $\mathbb{Q}$ is not.

If M_i are countably many topological spaces and for any i the set $U_i \subset M_i$ is open, we can define the **restricted topological product** of the spaces M_i with respect to U_i to be the set of all $m = (m_i) \in \prod_i M_i$ such that $m_i \in U_i$ for all but finitely many i. We can define a topology on the restricted topological product as follows: Each of the sets $\prod_i A_i$ (where A_i is open in M_i and $A_i = U_i$ for all but finitely many i) is open, and a subset B of the restricted topological product is open if and only if it is a countable union of sets of the above form. If the sets M_i and U_i carry additional structure, this is often passed on to the restricted topolological product. For example, the restricted topological product is locally compact if the spaces M_i are locally compact and the sets U_i are compact (see [Cas67]). If M_i and U_i are topological rings, then the restricted topological product is again a topological ring. We will see an example of this concept (where the restricted topological product is both locally compact and a ring) in Chapter I.4.

0.1.2 Differential, Riemannian and hyperbolic manifolds

Paracompact Haussdorff spaces that behave locally like the Euclidean space $\mathbb{R}^m$ are the main objects of study in differential geometry.

Definition 0.1.2 A **topological manifold** M of dimension m is a connected, paracompact Haussdorff space that is **locally Euclidean** (i.e., for any $p \in M$ there is a $U \subset M$ with $p \in U$, an open set $\Omega \subset \mathbb{R}^m$ and a homeomorphism $x : U \to \Omega$). Manifolds of dimension m are also called ***m*-manifolds**.

Note that we do not explicitely mention the topology of M, but we use it implicitely when speaking about homeomorphisms. Any such homeomorphism $x : U \to \Omega$ is called **chart**. An **atlas** $\mathcal{A}$ of M is a set of charts such that any point $p \in M$ lies in the domain of a chart $x \in \mathcal{A}$. An altas is called **differentiable** if for any two charts $x_1 : U_1 \to \Omega_1, x_2 : U_2 \to \Omega_2$ with $U_1 \cap U_2 \neq \emptyset$ the **transition map** $x_2 \circ x_1^{-1} : x_1(U_1 \cap U_2) \to x_2(U_1 \cap U_2)$ is differentiable. Let $\mathcal{A}$ be a differentiable atlas. A chart $x : U \to \Omega$ is called **compatible** with $\mathcal{A}$ if $\mathcal{A} \cup \{x\}$ is again a differentiable atlas. A differentiable atlas $\mathcal{A}$ is called a **differentiable structure** if all charts that are compatible to $\mathcal{A}$ belong to $\mathcal{A}$.

Definition 0.1.3 A topological manifold equipped with a differentiable structure is called a **differentiable manifold**.

We will mainly need a special kind of differentiable manifold:

Definition 0.1.4 A **Riemannian manifold** (M, g) is a differentiable manifold together with a **Riemannian metric** g on M.

We will not define what a Riemannian metric is (this would need the tangent space first and, for proper understanding, tensor fields on manifolds). Instead we will just mention how Riemannian metrics look like. We hope that this, together with Example 0.1.6, will suffice to understand Riemannian metrics enough for our purpose.

If M has dimension m, then a Riemannian metric g associates to each point $p \in M$ a positive definite symmetric bilinear form $g(p)$ on some m-dimensional vector space of differential operators, i.e., g is a differential form. Thus, for $m = 2$, a Riemannian metric can be written as

$$g = g_{1,1}\,\mathrm{d}x_1^2 + g_{1,2}\,\mathrm{d}x_1\,\mathrm{d}x_2 + g_{2,2}\,\mathrm{d}x_2^2$$

or, equivalently,

$$g(p) = \begin{pmatrix} g_{1,1}(p) & \frac{1}{2}g_{1,2}(p) \\ \frac{1}{2}g_{1,2}(p) & g_{2,2}(p) \end{pmatrix}.$$

If (M, g) is a Riemannian manifold, we can measure lengths and distances on M. Let $\gamma : [a, b] \to M$ be a smooth curve (we call a curve γ **smooth** if for any chart $x : U \to \Omega$ with suitable $U \subset M$ the curve $x \circ \gamma$ is smooth). Then the **length** of γ is defined as

$$L(\gamma) := \int_a^b \sqrt{g(\gamma(t))(\dot\gamma(t), \dot\gamma(t))}\,\mathrm{d}t,$$

where $\dot\gamma(t) = \frac{\mathrm{d}}{\mathrm{d}t}\gamma(t)$. The **distance** between two points p_1, p_2 on M is the infimum over all lengths of curves from p_1 to p_2. If a curve γ locally minimizes the distance (compare Example 0.1.6), then γ is called **geodesic**. Another definition of geodesics can be found in [Jos08].

Finally we define hyperbolic manifolds. We will just state the definition for special Riemannian manifolds of dimension 2.

Definition 0.1.5 Let (M, g) be a Riemannian manifold of dimension 2 with Riemannian metric $g = g_1\,\mathrm{d}x^2 + g_2\,\mathrm{d}y^2$ (i.e., $g_{1,2} = 0$). The **curvature** K of (M, g) is defined as

$$K = \frac{-1}{\sqrt{g_1 g_2}}\left(\frac{\partial}{\partial x}\left(\frac{1}{\sqrt{g_1}}\frac{\partial}{\partial x}\sqrt{g_2}\right) + \frac{\partial}{\partial y}\left(\frac{1}{\sqrt{g_2}}\frac{\partial}{\partial y}\sqrt{g_1}\right)\right).$$

(M, g) is called **hyperbolic** if $K = -1$.

The usual definition of curvature is not the above. In fact, there are different curvatures (such as sectional curvature, Gauß curvature), cf. [Jos08]. The definition of these in general cases requires more notation. The general definition allows one to define the curvature for arbitrary Riemannian metrics and this yields a definition of hyperbolic manifolds for arbitrary dimensions. Since we will only need the above formula this will suffice for us.

Example 0.1.6 Let us consider the sphere $\mathscr{S}^2 := \{s \in \mathbb{R}^3 : s_1^2 + s_2^2 + s_3^2 = 1\}$ together with the topology inherited from $\mathbb{R}^3$. Then $\mathscr{S}^2$ is clearly connected and the topology defines a Haussdorff space. We show that $\mathscr{S}^2$ is a manifold. For that, we need to show that $\mathscr{S}^2$ is paracompact and locally Euclidean. We will even show that $\mathscr{S}^2$ is a differentiable manifold by defining a finite open covering and a differentiable structure.

Let $N := (0,0,1)$ and $S := (0,0,-1)$. Then the sets $U_N := \mathscr{S}^2 \backslash \{S\}$ and $U_S := \mathscr{S}^2 \backslash \{N\}$ define a finite open covering of $\mathscr{S}^2$. Consider the **stereographic projections** $x_N : U_N \to \mathbb{R}^2, x_S : U_S \to \mathbb{R}^2$, given by

$$x_N(s_1,s_2,s_3) := \frac{1}{1+s_3}(s_1,s_2), \quad x_S(s_1,s_2,s_3) := \frac{1}{1-s_3}(s_1,s_2).$$

The stereographic projection x_N is shown in Figure 0.1.1.

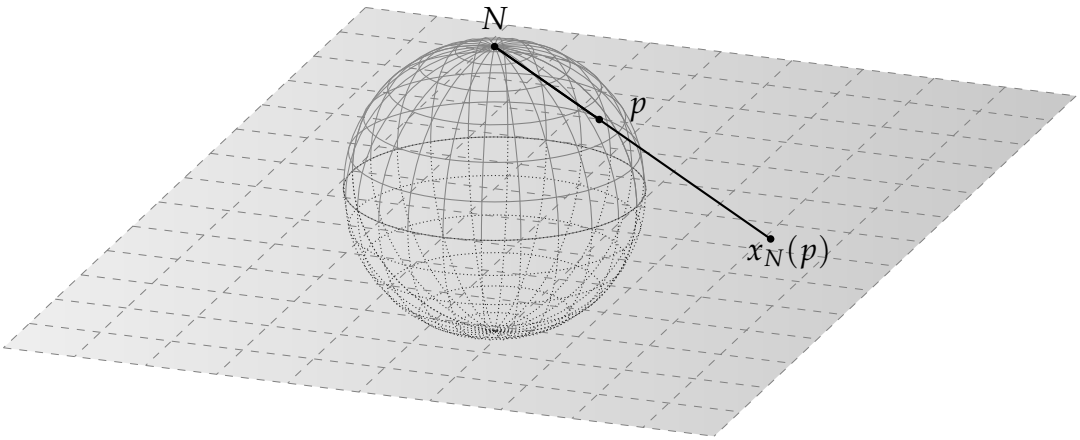

Figure 0.1.1: The stereographic projection x_N.

The stereographic projections are charts from $\mathscr{S}^2$ to $\mathbb{R}$. By computing the transition maps we see that $\{x_N, x_S\}$ is a differentiable atlas, thus we get a differentiable structure on $\mathscr{S}^2$ by taking the differentiable structure that contains this atlas.

We define a Riemannian metric g on S^2 by

$$g = \frac{4}{(1+u^2+v^2)^2}\left(\mathrm{d}u^2 + \mathrm{d}v^2\right), \text{ i.e., } g(p) = \begin{pmatrix} \frac{4}{(1+u^2+v^2)^2} & 0 \\ 0 & \frac{4}{(1+u^2+v^2)^2} \end{pmatrix},$$

where $(u, v) = x_N(p)$. We will measure the lengths of two curves and determine the curvature K.

Let $R := (0, 1, 0), F := (1, 0, 0) \in \mathscr{S}^2$. We consider two curves from R to F, namely

$$\gamma_1 : \left[0, \frac{\pi}{2}\right] \to \mathscr{S}^2, \quad \gamma_1(t) = (\sin t, \cos t, 0),$$
$$\gamma_2 : \left[0, \frac{3\pi}{2}\right] \to \mathscr{S}^2, \quad \gamma_2(t) = (-\sin t, \cos t, 0).$$

Note that since these curves lie in the $x-y-$plane, the stereographic projections of these curves are the identity maps. We get

$$L(\gamma_1) = \int_0^{\frac{\pi}{2}} \sqrt{\frac{4}{(1+(\sin t)^2+(\cos t)^2)^2}\left((\cos t)^2 + (\sin t)^2\right)}\,\mathrm{d}t = \int_0^{\frac{\pi}{2}} 1\,\mathrm{d}t = \frac{\pi}{2}$$

and similarly $L(\gamma_2) = \frac{3\pi}{2}$. These are exactly the values we expected since γ_1 is a quarter of the arc of a circle with radius 1 (analogous for γ_2). The curves γ_1 and γ_2 both locally minimize the length of curves connecting R and F (if we vary the curves by a small amount, the lengths increase), thus γ_1 and γ_2 are geodesics.

For the curvature K we compute

$$\frac{\partial}{\partial u}\left(\frac{1}{\sqrt{g_1}}\frac{\partial}{\partial u}\sqrt{g_2}\right) = \frac{1}{(1+u^2+v^2)^2}(4u^2 - 2(1+u^2+v^2)),$$

hence

$$K = \frac{-(1+u^2+v^2)^2}{4} \cdot \frac{4}{(1+u^2+v^2)^2} \cdot (u^2+v^2-(1+u^2+v^2)) = 1.$$

The sphere $\mathscr{S}^2$ will become important in another context. Consider the complex plane $\mathbb{C}$. We define the **Riemann sphere** to be $\widehat{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$. Now the stereographic projections define homeomorphisms from $\mathscr{S}^2$ to $\widehat{\mathbb{C}}$.

For more on manifolds see [Boo03, Jos08].

0.1.3 Hyperbolic geometry

Hyperbolic manifolds are important examples of **hyperbolic geometries**. We will see a detailed example in Paragraph 0.3.2. Hyperbolic geometry is a geometry satisfying the following axioms (compare [CFKP97]):

1. Two distinct points can be joined by exactly one straight line segment.
2. Any straight line segment can be extended in both directions arbitrarily long.
3. For any point P and any (positive) radius r there is exactly one circle with radius r whose center is P.
4. All right angles are congruent.
5. For any line l and any point P not on l there are at least two lines through P that do not intersect l.

Here a line is defined as the shortest curve between two points whereas a line in the context of Riemannian manifolds is defined to be a segment of a geodesic. The axiom that distinguishes hyperbolic geometry from Euclidean geometry is the fifth. In Euclidean geometry, given a line l and a point P not on l there is exactly one line through P that does not intersect l. This line is the parallel of l through P. In this style we call any line l' that does not intersect a given line l a **parallel** to l. Note that in hyperbolic geometry parallel lines are not equidistant.

There are some models for hyperbolic geometry, i.e., sets equipped with a metric that satisfies the hyperbolic axioms. Examples can be found in [RR94, CFKP97]. We will see one of those models in Paragraph 0.3.2.

The hyperbolic axioms have interesting consequences. Triangles have angle sum less than π and given three angles with sum less than π, there is (up to congruence) exactly one hyperbolic triangle with these angles. Moreover, the area of any triangle is bounded by π (compare Paragraph 0.3.2). More interesting properties and a comparison with other geometries can be found in [RR94, CFKP97].

0.1.4 Algebraic Curves

We conclude the basics about geometry with algebraic curves. Let us first recall the notion of affine plane and projective plane. We will not mention the exact definition of the affine plane but rather give some intuition.

Let K be a field (for visualization, take $K = \mathbb{R}$). Then we can view K^2 as **affine plane** $A^2(K)$ by considering elements of K^2 as **points**.

When we examine distinct lines in $A^2(\mathbb{R})$ they either intersect or they are parallel. We wish to expand the affine plane so that we do not need to distinguish between these two cases. This can be done by considering **points at infinity**. By defining them appropriately, parallel lines do intersect at infinity. This is the idea of the projective plane.

Definition 0.1.7 The **projective plane** $P^2(K)$ is the set of all lines in K^3 through the origin.

We denote elements in $P^2(K)$ by **homogeneous coordinates** $(x_0 : x_1 : x_2)$. By definition the equality $(x_0 : x_1 : x_2) = (y_0 : y_1 : y_2)$ holds if and only if there is a $\lambda \in K\backslash\{0\}$ such that $(x_0, x_1, x_2) = \lambda(y_0, y_1, y_2)$. The affine plane $A^2(K)$ can be canonically embedded in $P^2(K)$ by

$$\iota : A^2(K) \to P^2(K), \quad \iota(x_1, x_2) = (1 : x_1 : x_2).$$

Then the **points at infinity** are $P^2(K)\backslash\iota(A^2(K)) = \{(0 : x_1 : x_2) : x_1, x_2 \in K\}$.

We will consider curves defined on the affine or projective plane.

Definition 0.1.8 Let K be a field, $A^2(K)$ the affine plane over K and $P^2(K)$ the projective plane over K.

- A set $C \subset A(K^2)$ is called an **affine algebraic curve** if there is a polynomial $f \in K[x_1, x_2]$ with $\deg(f) \geq 1$ and

$$C = V(f) := \{(c_1, c_2) \in A^2(K) : f(c_1, c_2) = 0\}.$$

- A set $C \subset P(K^2)$ is called a **projective algebraic curve** if there is a homogeneous polynomial $f \in K[x_0, x_1, x_2]$ with $\deg(f) \geq 1$ and

$$C = V(f) := \{(c_0 : c_1 : c_2) \in P^2(K) : f(c_0, c_1, c_2) = 0\}.$$

Since there is a 1-to-1 correspondence

$$\{\text{polynomials } f \in K[x_1, x_2]\} \leftrightarrow \{\text{homogeneous polynomials } g \in K[x_0, x_1, x_2]\}$$

$$\sum a_{j,k} x_1^j x_2^k \to \sum a_{j,k} x_0^{\deg(f)-j-k} x_1^j x_2^k$$

$$\sum a_{j,k,l} x_1^j x_2^k \leftarrow \sum a_{j,k,l} x_0^l x_1^j x_2^k$$

we can restrict ourselves (at least when dealing with the defining polynomials) to affine algebraic curves and view projective algebraic curves as affine algebraic curves with some points at infinity.

If $C = V(f)$ is an affine algebraic curve and $f = f_1^{v_1} \cdots f_r^{v_r}$ is the decomposition of f into irreducible factors, we call $\widetilde{f} := f_1 \cdots f_r$ the **minimal polynomial** of C. Note that $V(\widetilde{f}) = V(f) = C$. Moreover, the minimal polynomial is unique up to units. We can use the minimal polynomial to define certain invariants or properties of algebraic curves.

Let $C = V(f)$ be an affine algebraic curve and f be a minimal polynomial of C. Then the **degree** of C is defined as the degree of f. The curve C is called **smooth** if

$$\left(\frac{\partial f}{\partial x_1}(c), \frac{\partial f}{\partial x_2}(c)\right) \neq (0,0)$$

for all $c \in C$ (for general fields we can take the formal derivative instead of the analytic derivative of f). If C is not smooth, then C is called **singular**.

More about algebraic curves can be found in [Fis94]

0.2 Complex Analysis

In this section we examine basic definitions and results in complex analysis, in particular about holomorphic functions.

0.2.1 Exponential sums

Exponential sums are complex sums that arise in several (not only number theoretic) contexts.

Definition 0.2.1 Let $f : \mathbb{N} \to \mathbb{R}$ be a function. Then an **exponential sum** is of the form

$$\sum_{a<n\leq b} e^{2\pi i f(n)}$$

for some $a, b \in \mathbb{R} \cup \{\infty\}$.

In some contexts we use the notations $e(x) := e^{2\pi i x}$ and $e_q(x) := e^{\frac{2\pi i x}{q}}$, so that the above sum becomes $\sum_{a<n\leq b} e(f(n))$. Whenever we use this notation we will explicitely mention this. There is one notable case in which we do not use this notation: In the theory of Fourier series it is not common to use this notation, and we will stick to the usual notation.

We do not use much about exponential sums in this thesis (and everything we need will be done in the respective chapters). In many cases one needs to bound exponential sums. There are some methods and results that apply in general settings, cf. [GK91].

0.2.2 Holomorphic and meromorphic functions

Functions depending on one complex variable that are (almost everywhere) complex differentiable play an important role in analytic number theory. Usually, we denote the variable by $s = \sigma + it$ where $\sigma, t \in \mathbb{R}$.

All definitions and results in the remaining paragraphs of this section can be found in [FB06]. In this section G always denotes a domain, i.e., an open connected subset $G \subset \mathbb{C}$.

Definition 0.2.2 Let $f : G \to \mathbb{C}$ be a complex function and $s_0 \in G$. Then f is called **complex differentiable** at s_0 if its **derivative**

$$f'(s_0) := \lim_{s \to s_0} \frac{f(s) - f(s_0)}{s - s_0}$$

exists. The function f is called **holomorphic** in G if f is complex differentiable for all $s \in G$.

Complex differentiability is much stronger than real differentiability. Any holomorphic funcion is automatically **smooth** (i.e., for any $k \in \mathbb{N}$, f is k times complex differentiable) and **analytic** (i.e., f can be expanded in a power series). In fact these properties are equivalent. More equivalent formulations of holomorphic functions can be found in [FB06].

Many interesting complex functions arising from number theory are not everywhere complex differentiable, they are not even defined everywhere. **Singularities** are points s_0 where f is not defined. For example, the functions $f(s) = \frac{1}{s}$ and $g(s) = \log(s)$ (where log denotes the principal value of the complex logarithm) both have singularities at $s_0 = 0$. We are especially interested in singularities that are discrete:

Definition 0.2.3 Let $f : G \to \mathbb{C}$ holomorphic and $s_0 \notin G$. Then s_0 is called **isolated singularity** if there is an open set U containing s_0 such that $U \backslash \{s_0\} \subset G$.

Thus, if s_0 is an isolated singularity of $f : G \to \mathbb{C}$, there is a set $V \subset G$ such that $V \cup \{s_0\}$ is open. We will call such a set a **neighboorhoud of s_0 in G**. In

the above examples, 0 is an isolated singularity of $f(s) = \frac{1}{s}$ whereas it is not an isolated singularity of $g(s) = \log(s)$.

There are three possible types of isolated singularities. Let $f : G \to \mathbb{C}$ be holomorphic and s_0 be an isolated singularity. Then s_0 is called

- **removable singularity** if there is a holomorphic function $g : G \cup \{s_0\} \to \mathbb{C}$ such that the restriction of g to G is f,
- **pole** if for any $R > 0$ there is a neighbourhood U of s_0 in G with $|f(s)| > R$ for all $s \in U$,
- **essential singularity** if s_0 is neither removable nor a pole.

A function $f : G \to \mathbb{C}$ is called **meromorphic** on G if there is a discrete set $P(f) \subset G$ such that $f : G \backslash P(f) \to \mathbb{C}$ is holomorphic and any $s \in P(f)$ is a pole or a removable singularity of f.

If f is a non-vanishing meromorphic function on G and $s_0 \in G$, there is an $k \in \mathbb{Z}$ and a holomorphic function g on G such that $g(s_0) \neq 0$ and

$$f(s) = (s - s_0)^k g(s)$$

for all $s \in G \backslash \{s_0\}$. The number $\operatorname{ord}_{s_0} f := k$ is called the **order** of f at s_0. Hence $\operatorname{ord}_{s_0} f$ is positive if $f(s_0) = 0$, $\operatorname{ord}_{s_0} f = 0$ if f has no pole at s_0 and $f(s_0) \neq 0$, and $\operatorname{ord}_{s_0} f$ is negative if f has a pole at s_0. If f has a pole at s_0, the **order of the pole** of f at s_0 is defined as $-\operatorname{ord}_{s_0} f$. If the order of the pole is 1, we call the pole **simple**.

In many cases complex functions f are given by integrals or series. Most of these integrals and series will exist, respectively converge, in a half plane $\sigma > \sigma_0$ for some $\sigma_0 \in \mathbb{R}$. If this is the case, we wish to extend f to a bigger domain (if possible, all of $\mathbb{C}$). This is done with analytic continuation.

Definition 0.2.4 Let $f : G \to \mathbb{C}, \widetilde{f} : \widetilde{G} \to \mathbb{C}$ be holomorphic functions with $G \subset \widetilde{G}$ such that the restriction of $\widetilde{f}$ on G is f. Then we call $\widetilde{f}$ **holomorphic continuation** or **analytic continuation** of f.

According to the **identity theorem** any two holomorphic functions f and g that coincide on a nonempty open set U are identical, thus any two holomorphic continuations to the same domain $\widetilde{G}$ are identical.

In most cases our holomorphic functions do not have an analytic continuation to $\mathbb{C}$, but a continuation to $\mathbb{C}$ such that the continuation is a meromorphic function. Then it is common to call this an analytic continuation besides some poles

at certain points. If we are not interested in the number and position of the poles, we will simply refer to a **meromorphic continuation**.

Often an analytic continuation is given via a **functional equation**. These are equations relating the value of a function at some point to the value at some other point. We will see this in Paragraph 0.5.4.

0.2.3 Laurent series

As already mentioned, holomorphic functions are analytic, i.e., they can be expanded in a power series. A similar statement holds for meromorphic functions.

Definition 0.2.5 A **Laurent series** is a symbol of the form

$$L(s) = \sum_{n=-\infty}^{\infty} a_n (s-s_0)^n := \sum_{n=1}^{\infty} \frac{a_{-n}}{(s-s_0)^n} + \sum_{n=0}^{\infty} a_n (s-s_0)^n,$$

where $a_i, s_0 \in \mathbb{C}$. We say that a Laurent series is **convergent** in $G \subset \mathbb{C}$ if both of the sums $\sum_{n=1}^{\infty} \frac{a_{-n}}{(s-s_0)^n}$ and $\sum_{n=0}^{\infty} a_n (s-s_0)^n$ are convergent in G.

If L is a Laurent series but no power series, then L either converges nowhere or in an open **annulus**, i.e., in a set of the form $\{s \in \mathbb{C} : r < |s-s_0| < R\}$ for some $r, R \in \mathbb{R}$ with $0 \leq r < R \leq \infty$.

If $f : G \to \mathbb{C}$ is holomorphic and $s_0 \notin G$ is an isolated singularity of f, then we can expand f in a Laurent series $f(s) = \sum_{n=-\infty}^{\infty} a_n (s-s_0)^n$ around s_0. The precise form depends on the type of the singularity:

- s_0 is removable if and only if the Laurent series is a power series,
- s_0 is a pole if and only if there is an $n_0 < 0$ such that $a_{n_0} \neq 0$ and $a_n = 0$ if $n < n_0$ (in fact we have $n_0 = \operatorname{ord}_{s_0} f$),
- s_0 is essential if and only if $a_{-n} \neq 0$ for infinitely many $n \in \mathbb{N}$.

If we expand f in a Laurent series, there is one coefficient that is of particular interest, especially if s_0 is a pole:

Definition 0.2.6 If $f : G \to \mathbb{C}$ is meromorphic and $f(s) = \sum_{n=-\infty}^{\infty} a_n (s-s_0)^n$ its Laurent series, then the coefficient a_{-1} is called the **residue** of f at s_0, denoted by $\operatorname{Res}(f; s_0)$.

0.2.4 Integral formulae

An important tool in complex analysis is the line integral.

Definition 0.2.7 Let $f : G \to \mathbb{C}$ be continuous and $\gamma : [a, b] \to \mathbb{C}$ be a smooth curve. The integral

$$\int_\gamma f(s)\,\mathrm{d}s := \int_a^b f(\gamma(\tau))\dot{\gamma}(\tau)\,\mathrm{d}\tau$$

is called the **line integral** of f alongside γ.

There are a lot of important integral formulae. We will need the following two:

Theorem 0.2.8 (Cauchy's integral formula) *Let $f : G \to \mathbb{C}$ be holomorphic and let $s_0 \in G$. Then*

$$f(s) = \frac{1}{2\pi i}\int_{C(s_0)} \frac{f(\tau)}{\tau - s}\,\mathrm{d}\tau$$

holds for any counter-clockwise oriented circle $C(s_0) \subset G$ with center at s_0 and for any s in the interior of $C(s_0)$.

Theorem 0.2.9 (Residue theorem) *Let $\gamma : [a, b] \to G$ be a simple closed curve (i.e., $\gamma(a) = \gamma(b)$ and $\gamma(r) \neq \gamma(s)$ for any other choices of $r, s \in [a, b]$). Let $s_1, \ldots, s_n$ be distinct points in the bounded open subset of G whose boundary is the trace of γ. If $f : G \to \mathbb{C}$ is a meromorphic function such that $f : G\backslash\{s_1, \ldots, s_n\} \to \mathbb{C}$ is holomorphic, we have*

$$\int_\gamma f(s)\,\mathrm{d}s = 2\pi i \sum_{i=1}^{n} \mathrm{Res}(f; s_i).$$

0.2.5 Möbius transformations

Now we turn our attention to special holomorphic functions.

Definition 0.2.10 Let $G_1, G_2 \subset \mathbb{C}$ be domains. A function $f : G_1 \to G_2$ is called **conformal** if f is bijective and holomorphic.

There is an equivalent geometric notion of conformal functions. When viewing the function involved as a mapping $f : \mathbb{R}^2 \to \mathbb{R}^2$, the conformal functions are exactly the functions that preserve angles and orientation, cf. [FB06].

If $f : G \to G$ is conformal, we call f an **automorphism**. The group of automorphisms of G (with composition) is denoted by $\mathrm{Aut}(G)$. We dicuss some special conformal mappings.

Definition 0.2.11 A **fractional linear transformation** f is a meromorphic function of the form

$$f(s) = \frac{as + b}{cs + d}$$

with $a, b, c, d \in \mathbb{C}$ and $ad - bc \neq 0$.

We wish to turn fractional linear transformations into holomorphic functions. Therefore we need to allow ∞ as argument and value of such a function. We do this by continuing a fractional linear transformation to the Riemann sphere $\widehat{\mathbb{C}}$ (recall Example 0.1.6). Any fractional linear transformation f can be uniquely extended to $\widehat{\mathbb{C}}$ via $f(\infty) = \infty$ if $c = 0$, or $f\left(-\frac{d}{c}\right) = \infty$ and $f(\infty) = \frac{a}{c}$ if $c \neq 0$. We call continuations of fractional linear transformations **Möbius transformations**. This continuation is a holomorphic function when viewed as a function from $\widehat{\mathbb{C}}$ to $\widehat{\mathbb{C}}$ (we will not treat holomorphic functions on $\widehat{\mathbb{C}}$ here).

In fact (when extending the notion of conformal mappings appropriately) the group of Möbius transformations is exactly the group of automorphisms of the Riemann sphere. We can describe Möbius transformations with matrices: The map

$$\phi : \mathrm{GL}_2(\mathbb{C}) \to \mathrm{Aut}\,\widehat{\mathbb{C}}, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \frac{az+b}{cz+d}$$

is a homomorphism of groups, its kernel is $\{\lambda I_2 : \lambda \in \mathbb{C}\}$.

0.3 The upper half plane $\mathbb{H}$

The upper half plane

$$\mathbb{H} := \{s \in \mathbb{C} : \Im(s) > 0\}$$

will become important in Chapter I.9. We will consider $\mathbb{H}$ in two ways: First as subset of $\mathbb{C}$ and second as a hyperbolic manifold.

0.3.1 ... as complex domain

Since $\mathbb{H} \subset \mathbb{C}$ is a domain we can examine the group of automorphisms of $\mathbb{H}$.

Theorem 0.3.1 *The group of automorphisms of $\mathbb{H}$ is given by*

$$\mathrm{Aut}(\mathbb{H}) = \left\{ \frac{az+b}{cz+d} : a, b, c, d \in \mathbb{R}, ad - bc > 0 \right\}.$$

In other words, the group of automorphisms of $\mathbb{H}$ is a subgroup of $\mathrm{Aut}(\widehat{\mathbb{C}})$, when identifying fractional linear transformations with their corresponding Möbius transformations. This is common, even without mentioning this identification. To make this exact, we could consider $\mathbb{H}$ together with its boundary (see Paragraph 0.3.2). These details are not important for us, therefore we will call automorphisms of $\mathbb{H}$ Möbius transformations.

When identifying these Möbius transformations with matrices again, we have a homomorphism

$$\phi : \mathrm{GL}_2^+(\mathbb{R}) \to \mathrm{Aut}(\mathbb{H}), \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \frac{az+b}{cz+d},$$

where $\mathrm{GL}_2^+(\mathbb{R})$ denotes all real 2×2 matrices with positive determinant. The kernel of this map is $\{\lambda I_2 : \lambda \in \mathbb{R}_{>0}\}$.

0.3.2 ... as hyperbolic manifold

Since $\mathbb{H}$ is an open subset of $\mathbb{C} \cong \mathbb{R}^2$, it is a differentiable manifold. We introduce a Riemannian metric g on $\mathbb{H}$ and investigate some properties of $(\mathbb{H}, g)$.

Let $z = x + iy$. Then we consider the Riemannian metric

$$g = \frac{\mathrm{d}x^2 + \mathrm{d}y^2}{y^2}, \text{ i.e., } g(x,y) = \begin{pmatrix} \frac{1}{y^2} & 0 \\ 0 & \frac{1}{y^2} \end{pmatrix}.$$

First we compute the curvature K, cf. [Ser13]. Since

$$\frac{\partial}{\partial x}\left(\frac{1}{\sqrt{g_1}}\frac{\partial}{\partial x}\sqrt{g_2}\right) = 0 \text{ and } \frac{\partial}{\partial y}\left(\frac{1}{\sqrt{g_1}}\frac{\partial}{\partial y}\sqrt{g_2}\right) = \frac{1}{y^2},$$

we get

$$K = \frac{-1}{\frac{1}{y^2}} \cdot \frac{1}{y^2} = -1,$$

hence $(\mathbb{H}, g)$ is a hyperbolic manifold.

Now we determine the geodesics on $\mathbb{H}$. There is a general equation that geodesics have to satisfy, see [Hit]. We will just look at our special case. In this case, for a curve $\gamma(t) = x(t) + iy(t)$ where $x(t)$ and $y(t)$ are the real and imaginary parts, respectively, the equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{x'(t)}{y(t)^2}\right) = 0 \text{ and } \frac{x'(t)^2 + y'(t)^2}{y(t)^2} = 1$$

need to hold. From the first equation we get $x'(t) = cy(t)^2$ for some $c \in \mathbb{R}$. Thus we have two possibilities.

Either $c = 0$, then $x(t)$ is constant and γ is a vertical line. In the other case, $c \neq 0$ and the two equations above give

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\frac{\mathrm{d}y}{\mathrm{d}t}}{\frac{\mathrm{d}x}{\mathrm{d}t}} = \frac{\sqrt{y^2 - c^2y^4}}{cy^2} = \frac{\sqrt{1 - c^2y^2}}{cy}.$$

Then we get $\frac{cy\,\mathrm{d}y}{\sqrt{1-c^2y^2}} = \mathrm{d}x$ and integrating this with respect to t results in

$$-\frac{1}{c}\sqrt{1 - c^2y^2} = x - a$$

for some $a \in \mathbb{R}$. Thus, γ satisfies the equation $(x-a)^2 + y^2 = \frac{1}{c^2}$, i.e., it is a semicircle with center on the real line.

This characterization immediately yields that for any two distinct points z_1, z_2 in the upper half plane $\mathbb{H}$ there is exactly one geodesic through z_1 and z_2 (in contrast to the sphere in Example 0.1.6).

Thus, to get the distance between two points we only need to compute the length of the geodesic segment between z_1 and z_2. There are a lot of (of course equivalent) formulae for the **hyperbolic distance** $d_{\mathbb{H}}$, see [Hit, Ser13, Kat10]. One of them is

$$d_{\mathbb{H}}(z_1, z_2) = 2\tanh^{-1}\left(\left|\frac{z_2 - z_1}{z_2 - \overline{z_1}}\right|\right).$$

In particular we have $d_{\mathbb{H}}(x + ai, x + bi) = \log\frac{b}{a}$ for $b > a > 0, x \in \mathbb{R}$ (we would also get this formula easily with the definition of the length, cf. Paragraph 0.1.2), i.e., the distance of two points on the same vertical line does not depend on the real part and increases with decreasing imaginary part. The situation is different for the distance of points on a horizontal line. Let $z_1 = x_1 + ti$ and $z_2 = x_2 + ti$. Then

$$d_{\mathbb{H}}(z_1, z_2) = 2\tanh^{-1}\left(\frac{|x_2 - x_1|}{\sqrt{(x_2 - x_1)^2 + 4t^2}}\right).$$

For constant x_1, x_2 and increasing t, this goes to 0.

Another direct consequence of the characterization of geodesics is that for any given line l there are infinitely many parallels to l, cf. Figure 0.3.1. Thus we are indeed dealing with a hyperbolic geometry.

In the characterization of the geodesics we see that the points on the real line play an important role. Note that for any given point $z = x + iy \in \mathbb{H}$, the hyperbolic distance from z to the real line is

$$\lim_{t\to 0} d_{\mathbb{H}}(x + iy, x + tiy) = \lim_{t\to 0} \log \frac{1}{t} = \infty.$$

Thus the real line can be seen as infinitely far away. As we have already seen, the distance of two points $z_1 = x_1 + ti, z_2 = x_2 + ti$ goes to 0 for increasing t, i.e., any sequence $z_n = x_n + it_n$ "diverges to the same point" when $t_n \to \infty$. Denote this point by $\infty_{\mathbb{H}}$. Then we call $\partial\mathbb{H} = \mathbb{R} \cup \infty_{\mathbb{H}}$ the **boundary at infinity** of $\mathbb{H}$. Points of $\partial\mathbb{H}$ are called **ideal points**.

As already statet in Paragraph 0.1.3, the angles and the area of triangles satisfy interesting properties. In our case we can verify these with the help of a formula from differential geometry, the **Gauß-Bonnet formula**. We will not explain all notation of this formula, for more information see [Lee97].

Theorem 0.3.2 (Gauß-Bonnet formula) *Let (M, g) be a Riemannian surface and let $D \subset M$ be a connected subset with piecewise differentiable boundary ∂D. Let $\chi(D)$ denote the Euler characteristic of D, K the curvature of M and κ_g the geodesic curvature along ∂D. Then*

$$\int_D K \,\mathrm{d}A + \int_{\partial D} \kappa_g(s) \,\mathrm{d}s = 2\pi\chi(D).$$

Let D be a triangle in $\mathbb{H}$. Then we have (compare [Ser13, Kat10]) $K = -1$ and $\kappa_g(s) = 0$ at each point where ∂D is differentiable (since the boundary consists of geodesic segments). If α_i denote the interior angles on the corners of ∂D (i.e., the points where ∂D is not differentiable), we have

$$\int_{\partial D} \kappa_g(s) \,\mathrm{d}s = \sum_i (\pi - \alpha_i).$$

Since $\chi(D) = 1$, we get

$$\int_D -1 \,\mathrm{d}A + \sum_i (\pi - \alpha_i) = 2\pi$$

and hence

$$A = \int_D 1 \,\mathrm{d}A = \pi - \alpha_1 - \alpha_2 - \alpha_3.$$

An immediate consequence is that both the angle sum in and the area of an hyperbolic triangle are less than π.

Strictly speaking, we have not defined what an hyperbolic angle of two crossing geodesics is. We will not do this here (see [Kat10] for a precise definition involving the usual formula with scalar products). We just mention that this definition will result in the same angles as if we would compute the Euclidean angles of the tangents of the geodesics, cf. [Kat10].

Example 0.3.3 Consider $z_1 = 1, z_2 = 3+2i, z_3 = 3+(2+\sqrt{8})i$ (i.e., $z_2, z_3 \in \mathbb{H}$ and $z_1 \in \partial\mathbb{H}$ is an ideal point). The geodesic going through z_1 and z_2 is the semicircle with center 3 and radius 2, the geodesic going through z_1 and z_3 is the semicircle with center $5+\sqrt{8}$ and radius $4+\sqrt{8}$, and the geodesic going through z_2 and z_3 is the vertical line $x = 3$ (see Figure 0.3.1). The triangle D defined by these three points has interior angles $\alpha_{z_1} = 0, \alpha_{z_2} = \frac{\pi}{2}$ and $\alpha_{z_3} = \frac{\pi}{4}$, thus D has hyperbolic area $\frac{\pi}{4}$. We further have

$$d_{\mathbb{H}}(z_2, z_3) = \tanh^{-1}(\sqrt{2}-1) = \frac{1}{2}\log(\sqrt{2}+1) \approx 0.4406868.$$

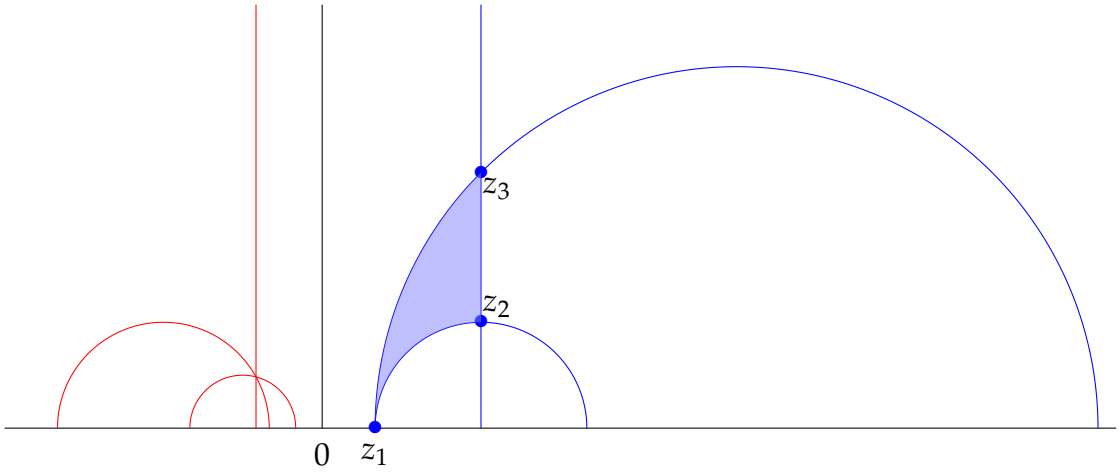

Figure 0.3.1: A hyperbolic triangle and parallel lines.

Any of the red geodesics in Figure 0.3.1 is a parallel of any blue geodesic, and in this style one can construct infinitely many lines through a given point parallel to a given line.

An important concept in geometry are isometries:

Definition 0.3.4 A map $f : \mathbb{H} \to \mathbb{H}$ is called **isometry** if f is differentiable (as function $f : \mathbb{R}^2 \to \mathbb{R}^2$) and $d_{\mathbb{H}}(f(z_1), f(z_2)) = d_{\mathbb{H}}(z_1, z_2)$ for any $z_1, z_2 \in \mathbb{H}$.

We have already seen isometries of $\mathbb{H}$: Every automorphism of $\mathbb{H}$ is also an isometry of $\mathbb{H}$. These are all isometries that preserve orientation. The group of all isometries of $\mathbb{H}$ is genereated by $\mathrm{Aut}(\mathbb{H})$ and the reflection $z \mapsto \bar{z}$.

Isometries are important because they preserve geometric objects and properties such as geodesics, lengths, angles, areas, triangles and much more. In Chapter I.9, we will be concerned with functions on $\mathbb{H}$ that allow special transformations under actions of some subsets of the group of isometries.

A more detailed discussion about isometries of $\mathbb{H}$ can be found in [Kat10, Ser13]. More about the hyperbolic plane and trigonometry on $\mathbb{H}$ can be found in [Wal16, Ser13].

0.4 Linear Algebra and Graph Theory

Now we examine some basics about special subsets of $\mathbb{R}^n$ as well as some linear algebra and graph theory.

0.4.1 Lattices and subsets of $\mathbb{R}^n$

Definition 0.4.1 Let $n \in \mathbb{N}$. A **lattice** Γ is a discrete subgroup of $\mathbb{R}^n$, i.e., a subgroup of the form

$$\Gamma = \mathbb{Z}v_1 + \cdots + \mathbb{Z}v_m$$

with $m \leq n$ and linearly independent vectors $v_i \in \mathbb{R}^n$. A lattice is called **full** if $m = n$.

In this thesis all lattices are assumed to be full. If Γ is a lattice, we call any set of the form $\gamma + P$, where $\gamma \in \Gamma$ and P is the parallelotope generated by $v_1, \ldots, v_n$, a **fundamental domain** of Γ and we define the volume of Γ to be

$$\mathrm{vol}(\Gamma) := \mathrm{vol}(P) = |\det(v_1, \ldots, v_n)| .$$

Example 0.4.2 Figure 0.4.1 shows the lattice

$$\Gamma = \left\{ k_1 \begin{pmatrix} 2 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} : k_1, k_2 \in \mathbb{Z} \right\}$$

and two possible fundamental domains. The volume of Γ is

$$\mathrm{vol}(\Gamma) = \left| \det \left(\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \right) \right| = 4.$$

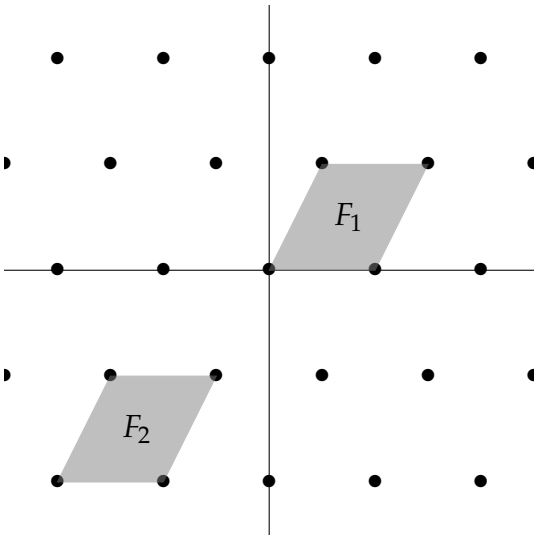

Figure 0.4.1: A lattice Γ and two fundamental domains F_1 and F_2.

Fundamental domains arise in other mathematical contexts, too. The general concept deals with group actions. Recall that a **group action** of a group $(G, \circ)$ with neutral element e on a set M is a map $* : G \times M \to M$ such that for all $g, h \in G$ and $m \in M$ the following two conditions hold:

- $e * m = m$,
- $(g \circ h) * m = g * (h * m)$.

Two elements $m_1, m_2 \in M$ are called **equivalent** under G if there is a $g \in G$ such that $g * m_1 = m_2$.

Definition 0.4.3 Let X be a topological space and G a group acting on X. An open subset $F \subset X$ is called **fundamental domain** (under the action of G) if the following two conditions hold:

- If $x, y \in F$ are distinct, there is no $g \in G$ with $g(x) = y$.
- For any $x \in X$ there is a $y \in \overline{F}$ and a $g \in G$ with $g(x) = y$.

Thus in this general setting a fundamental domain is a subset $F \subset X$ such that no distinct two elements in F are equivalent under G, but any element in X is equivalent to some y in the closure of F. There is also a more precise definition of fundamental domains where F is not required to be open and the element y in the second condition does not only belong to $\overline{F}$, but to F itself. In this case it is harder to determine a fundamental domain, thus we will stick to the above definition. Any fundamental domain of a lattice is a fundamental domain of the group action defined via the translations by the vectors v_i.

In any setting, fundamental domains contain all the information that we need, and in most interesting cases these fundamental domains are (at least partially) bounded, thus it is much easier to work with them.

When working with lattices, we also need two special kinds of subsets of $\mathbb{R}^n$.

Definition 0.4.4 Let $S \subset \mathbb{R}^n$ be a subset.

- S is called **centrally symmetric** if for any $\mathbf{x} \in S$ we have $-\mathbf{x} \in S$.
- S is called **convex** if for any $\mathbf{x}, \mathbf{y} \in S$ the line connecting $\mathbf{x}$ and $\mathbf{y}$ lies in S.

0.4.2 Polynomials and formal Laurent series

Definition 0.4.5 Let $n \in \mathbb{N}$ and R be a ring. We call a polynomial $f \in R[x_1, \ldots, x_n]$ **homogeneous of degree k** if $f(\lambda \mathbf{x}) = \lambda^k f(\mathbf{x})$ for any $\mathbf{x} \in R^n$ and $\lambda \in R$. Homogeneous polynomials are also called **forms**. If $n = 2$, the forms are called **binary**, if $n = 3$, they are called **ternary**. Forms of degree 1 are called **linear**, those of degree 2 are called **quadratic**.

There is a well developed theory about quadratic forms (in particular binary and ternary), but we do not need any results here.

If $f = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in R[x]$ is a **monic** polynomial (i.e., a polynomial with leading coefficient 1) in one variable with roots α_i (where the roots lie in a suitable extension of R and are counted with multiplicities), the **discriminant** of f is defined as

$$\Delta_f = \prod_{i<j} (\alpha_i - \alpha_j)^2 = (-1)^{\frac{n(n-1)}{2}} \prod_{i \neq j} (\alpha_i - \alpha_j).$$

A canonical generalization of polynomials are formal power series and formal Laurent series.

Definition 0.4.6 Let R be a ring.

- A symbol of the form $\sum_{n=0}^{\infty} a_n x^n$ with $a_n \in R$ and an indeterminate x is called **formal power series**. The set of formal power series is denoted by $R[[x]]$.
- A symbol of the form $\sum_{n=k}^{\infty} a_n x^n$ (for some $k \in \mathbb{Z}$) with $a_n \in R$ and an indeterminate x is called **formal Laurent series**. The set of formal Laurent series is denoted by $R((x))$.

With the usual operations, $R[[x]]$ is a ring and $R((x))$ is a field. If $f \in R((x))$ is any formal Laurent series, we can formally differentiate f with the usual rule for powers, i.e., if $f = \sum_{n=k}^{\infty} a_n x^n$, then its **formal derivative** f' is given by

$$f' = \sum_{n=k}^{\infty} n a_n x^{n-1}.$$

If $f = \sum_{n=0}^{\infty} f_n x^n \in R[[x]]$ and $g = \sum_{n=0}^{\infty} g_n x^n \in R[[x]]$ are two power series, their **Hadamard product** is defined to be

$$f \odot g = \sum_{n=0} (f_n \cdot g_n) x^n.$$

0.4.3 Matrix operations

We will need some special matrix operations.

The **Kronecker product** of two matrices A and B is the matrix

$$A \otimes B = \begin{pmatrix} a_{1,1}B & \cdots & a_{1,n}B \\ \vdots & \ddots & \vdots \\ a_{m,1}B & \cdots & a_{m,n}B \end{pmatrix}.$$

If A is an $m \times n$ matrix and B is an $r \times s$ matrix, then $A \otimes B$ is an $mr \times ns$ matrix. Note that in general $A \otimes B \neq B \otimes A$.

There is a relation between the Kronecker product and the matrix product. In fact, we have

$$(A \otimes B)(C \otimes D) = AC \otimes BD$$

whenever the respective products are defined. In particular, if A and B commute, then $A \otimes B$ and $B \otimes A$ commute since

$$(A \otimes B)(B \otimes A) = AB \otimes BA = BA \otimes AB = (B \otimes A)(A \otimes B).$$

If λ_j and μ_l are the eigenvalues of A, respectively B, (with multiplicities) with eigenvectors v_j, respectively w_l, then the eigenvalues of $A \otimes B$ are $\lambda_j \mu_l$ with eigenvectors $v_j \otimes w_l$. The above results and more about the Kronecker product can be found in [HJ08].

Related to the Kronecker product is the **Kronecker sum**. If A is an $m \times m$ matrix and B is an $n \times n$ matrix, then the Kronecker sum of A and B is defined as $A \oplus B = A \otimes I_n + I_m \otimes B$.

As usual, one can define the Kronecker sum and the Kronecker product of linear maps via the respective operation on representation matrices.

We will need one more result about the eigenvalues of sums of matrices. Suppose A and B commute. Then A and B can be simultaneously diagonalized, i.e., there is a matrix S such that $S^{-1}AS = D_A$ and $S^{-1}BS = D_B$ are diagonal matrices. Then $S^{-1}(A+B)S = S^{-1}AS + S^{-1}BS = D_A + D_B$. Thus the eigenvalues ν_k of $A+B$ are of the form $\lambda_j + \mu_l$, where λ_j and μ_l are eigenvalues of A respectively B belonging to the same eigenvector.

0.4.4 The Smith normal form

Similarly to matrices over fields there are normal forms of matrices with entries in a principal ideal domain. We will use the Smith normal form.

Definition 0.4.7 A matrix A (with entries in some principal ideal domain R) is said to be in **Smith normal form** if there are $a_i \in R$ such that $a_i | a_{i+1}$ and

$$A = \begin{pmatrix} a_1 & & & & & \\ & \ddots & & & & \\ & & a_m & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

Note that A does not need to be a square matrix. Any matrix over a principal ideal domain can be brought in Smith normal form, see [Mor05].

Theorem 0.4.8 *Let R be a prinicipal ideal domain and A be a matrix with entries in R. Then there are invertible matrices P, Q such that PAQ is in Smith normal form.*

0.4.5 Graphs

Now we lay the foundation in graph theory. For more about graphs see [BM08].

Definition 0.4.9 An **(undirected) graph** G is an ordered pair $G = (V, E)$ where V is a finite set and E is a subset of $\binom{V}{2} := \{\{v, w\} : v, w \in V, v \neq w\}$. We call the elements of V **vertices** and the elements of E **edges**.

Vertices v, w such that $\{v, w\} \in E$ are called **adjacent**. Sometimes one allows a vertex to be adjacent to itself. Edges from a vertex v to v are called **loops**. Unless stated otherwise, all graphs are assumed to be loopless.

One can also specify a direction of edges in a graph: A **directed graph** G is an ordered pair $G = (V, E)$ where V is a finite set and E is a subset of $V \times V$. Often one writes $D = (V, A)$ for directed graphs to distinguish between undirected and directed graphs. In the case of directed graphs edges are also called **arcs**.

Graphs exist as abstract objects. To illustrate them we usually draw graphs in the plane $\mathbb{R}^2$. Then vertices will be drawn as dots or small circles and edges as lines or curves. If G is a directed graph, the direction of the arcs is indicated with arrows. In general, we need to distinguish between a graph and its drawing (in particular when we are concerned about the relationship of edges in the plane, such as crossings). For our purposes this is not important, thus we will call (by abuse of notation) drawings of a graph again a graph. Figure 0.4.2 shows an (undirected) graph, the so called **Petersen graph**.

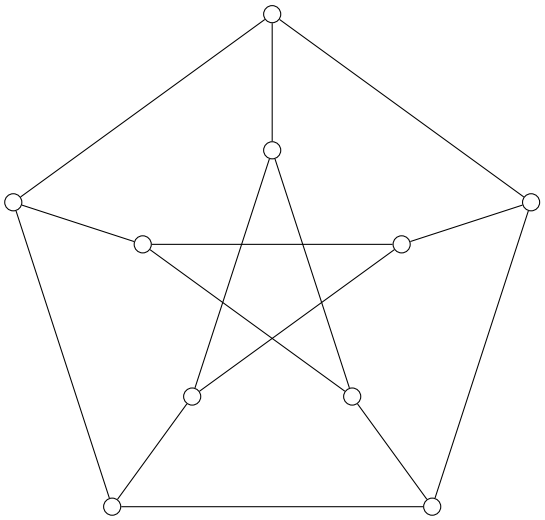

Figure 0.4.2: The Petersen graph PET.

In this thesis, when referring to a graph we will always mean an undirected graph. Directed graphs are always denoted as such. Since we will mostly be concerned with undirected graphs, we will define all relevant notations only for this case. For most concepts there are obvious analogues for directed graphs, cf. [BM08].

First we examine the relationships between vertices.

Definition 0.4.10 Let $G = (V, E)$ be a graph.

- G is called **complete** if $E = \binom{V}{2}$, i.e., if any two different vertices are adjacent.
- The **complement** of G is the graph $G^c := (V, E^c)$ with $E^c := \binom{V}{2} \setminus E$.

If G is a graph, we are also interested in certain concepts of reachability.

Definition 0.4.11 Let $G = (V, E)$ be a graph.

- A **path** in G is a sequence $P = v_0, \ldots, v_k$ of vertices such that $v_i \neq v_j$ for $i \neq j$, and $\{v_{i-1}, v_i\} \in E$ for $i = 1, \ldots, k$. The **length** of P is $l(P) := k \in \mathbb{N}_0$.
- G is called **connected** if for any $v, w \in V$ there is a path connecting v and w.
- The **distance** between two vertices $v, w \in V$ is

$$d(v, w) := \min\{l(P) : P \text{ is a path that connects } v \text{ and } w\}.$$

- The **diameter** of a graph is defined as $\operatorname{diam}(G) := \max_{v,w \in V} d(v, w)$.
- The ***k*-neighbourhood** of v is defined as $N_k(v) := \{w \in V : d(v, w) = k\}$.

Of special interest is the set of all vertices that are adjacent to a given vertex. The **neighbourhood** of a vertex v is the set $N(v) := \{w \in V : \{v, w\} \in E\}$. Note that $N(v) = N_1(v)$. The **degree** of a vertex v, denoted $d(v)$, is defined as the number of neighbours of v, i.e., $d(v) := |N(v)|$. A graph is called ***k*-regular** if all of its vertices have degree k.

A graph has a lot of different characteristic numbers. We will mention three of them: The clique number, the chromatic number and the vertex connectivity.

A **clique** is a subset $W \subset V$ such that any two different vertices of W are adjacent. A clique of size 3 is called **triangle**. The **clique number** $\omega(G)$ is the size (i.e., the number of vertices) of a maximal clique in G.

An **independent set** is a subset $W \subset V$ such that no two different vertices of W are adjacent. A graph $G = (V, E)$ is called ***k*-partite** if V is a disjoint union $V = V_1 \cup \cdots \cup V_k$ such that any V_i is an independent set. The smallest number k such that G is k-partite is called the **chromatic number** $\chi(G)$.

A graph G is called ***k*-connected** (with $k \geq 1$) if G has at least $k + 1$ vertices and G remains connected whenever at most $k - 1$ vertices of G are removed. The largest k such that G is k-connected is called **vertex connectivity** of G, denoted $\kappa(G)$.

We conclude this paragraph with some special graphs, automorphisms, and an example for the concepts defined here.

Definition 0.4.12 Let $(H, \circ)$ be a finite group with neutral element e_H and $S \subset H$ a subset with $e_H \notin S$ and $S^{-1} := \{s^{-1} : s \in S\} = S$. The **Cayley graph** $X(H, S)$ is the graph $G = (V, E)$ with $V = H$ and $\{v, w\} \in E :\Leftrightarrow v \circ w^{-1} \in S$.

Like for many other objects, one can define homomorphisms of graphs. Let $G_1 = (V_1, E_1)$ and $G_1 = (V_2, E_2)$ be graphs. A map $\varphi : V_1 \to V_2$ is called **isomorphism** if φ is bijective and for all $v, w \in V_1$ we have $\{v, w\} \in E_1$ if and only if $\{\varphi(v), \varphi(w)\} \in E_2$, i.e., two vertices are adjacent if and only if their respective images are adjacent. As usual one writes $G_1 \cong G_2$ if there is a isomorphism between G_1 and G_2. An isomorphism from G to G is called **automorphism** of G. The set $\mathrm{Aut}(G)$ of all automorphisms together with composition form the **group of automorphisms** of G. G is called **transitive** if $\mathrm{Aut}(G)$ acts transitively on V, i.e., if for any $v, w \in G$ there is a $\varphi \in \mathrm{Aut}(G)$ with $\varphi(v) = w$. It is easy to show that Cayley graphs are transitive, see [GR01].

A graph G is called **circulant** if $G \cong X(\mathbb{Z}/n\mathbb{Z}, S)$ for some $n \in \mathbb{N}, S \subset \mathbb{Z}/n\mathbb{Z}$.

Definition 0.4.13 Let $G = (V, E)$ be a graph and $V = \{v_1, \ldots, v_n\}$. Define a matrix $A = (a_{i,j})$ by

$$a_{i,j} = \begin{cases} 1, & \{v_i, v_j\} \in E \\ 0, & \{v_i, v_j\} \notin E \end{cases}.$$

The matrix A is called **adjacency matrix** of G. Since A is real and symmetric its eigenvalues are real. A graph is called **integral** if all eigenvalues of A are actually integers.

If A_1, A_2 are two adjacency matrices corresponding to different numberings of the vertices of a graph G, then there is a permutation matrix P such that $P^{-1}A_1P = A_2$. Hence the integrality of a graph does not depend on the numbering of its vertices.

Example 0.4.14 The Petersen graph is a connected, 3-regular graph, there are only cliques of size 1 and 2, thus $\omega(\mathrm{PET}) = 2$ and PET has no triangles. The coloring of the edges in Figure 0.4.3 define independent sets and thus a 3-partition of the Petersen graph. It is easy to show that the Petersen graph is neither 1- nor 2-partite, so $\chi(\mathrm{PET}) = 3$. The diameter is $\mathrm{diam}(\mathrm{PET}) = 2$, the red path in Figure 0.4.3 is the shortest path connecting v and w. The Petersen graph is 3-connected but not 4-connected (since deleting the three neighbours of any vertex would result in a graph that is not connected), hence $\kappa(\mathrm{PET}) = 3$. It can be shown that PET is not a Cayley graph, see [GR01].

From our drawing of the Petersen graph it is obvious that any automorphism of the regular 5-gon induces an automorphism of PET, in particular $\mathrm{Aut}(\mathrm{PET})$

acts transitively on both the outer and the inner vertices. Moreover, the automorphism depicted in Figure 0.4.4 maps inner vertices to outer vertices and vice versa. Combining these automorphisms shows that PET is transitive.

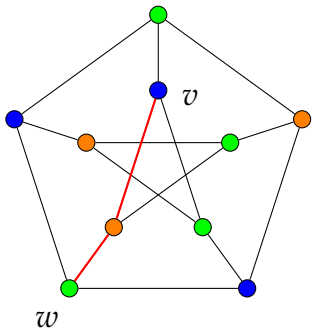

Figure 0.4.3: Properties of the Petersen graph PET.

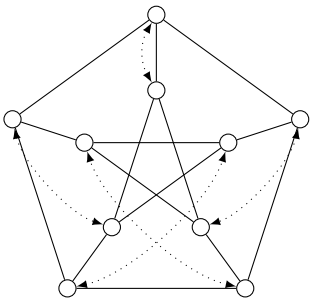

Figure 0.4.4: An automorphism of the Petersen graph PET.

The adjacency matrix (for some numbering of the vertices) of PET is

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}.$$

Its eigenvalues are -2 (with multiplicity 4), 1 (with multiplicity 5), and 3, thus the Petersen graph is integral.

0.5 Number Theory

We will assume that the reader is familiar with most number theoretic objects. Nevertheless we will recall some concepts and results here. While the concepts and definitions will be used (more or less) in later chapters, we will not use most of the results but instead see some possible proofs for these, using different techniques.

0.5.1 Miscellaneous

We will start to recall some basic notation and concepts.

If $n \in \mathbb{N}$ and $p \in \mathbb{P}$, we denote by $v_p(n)$ the exponent of p in the prime decomposition of n (here the v stands for valuation, cf. Paragraph 0.5.8), i.e., we have $n = \prod_{p\in\mathbb{P}} p^{v_p(n)}$. We will denote the **Euler totient function** with φ, i.e.,

$$\varphi(n) = \sum_{\substack{1\le k\le n \\ \gcd(k,n)=1}} 1 = \prod_{\substack{p\in\mathbb{P} \\ p|n}} (p-1)p^{v_p(n)-1}.$$

In this thesis the **number of divisors** of n will be denoted by $\tau(n)$, i.e.,

$$\tau(n) = \sum_{d|n} 1 = \prod_{\substack{p\in\mathbb{P} \\ p|n}} (v_p(n)+1).$$

The **divisor sum** of n is

$$\sigma(n) = \sum_{d|n} d = \prod_{\substack{p\in\mathbb{P} \\ p|n}} \frac{p^{v_p(n)+1}-1}{p-1}.$$

Recall that n is called **perfect** if $\sigma(n) = 2n$. The even perfect numbers are perfectly understood, see [SF07]:

Theorem 0.5.1 *An even number n is perfect if and only if it is of the form $2^{p-1}(2^p-1)$ such that both p and 2^p-1 are primes.*

Numbers of the form $2^n - 1$ are called **Mersenne numbers**. They can only be prime if n is prime, but this is not a sufficient condition, as $2^{11} - 1$ is not a prime. It is not known whether or not there are infinitely many Mersenne primes. For a list of known Mersenne primes see [MPS].

Until now, no odd perfect number is known, and it is believed that there are none. For some conditions that these numbers would have to fulfill (there are many more conditions derived in different articles), see [Bra43, Küh49, HM72].

Recall that for any $b \in \mathbb{N}_{\ge 2}$ there is a unique way to represent a natural number n in base b, i.e., the representation $n = \sum_{k=0}^{\infty} n_k b^k$ with $0 \le n_k < b$ is unique. We will denote the **base b representation** of n by $\langle n\rangle_b$, i.e., $\langle n\rangle_b = \ldots n_2 n_1 n_0$.

For $a, b \in \mathbb{Z}$, we denote by $\left(\frac{a}{b}\right)$ the **Kronecker symbol**, i.e.,

$$\left(\frac{a}{1}\right) = 1, \qquad \left(\frac{a}{0}\right) = \begin{cases} 1, & a = \pm 1 \\ 0, & a \neq \pm 1 \end{cases}, \qquad \left(\frac{a}{-1}\right) = \begin{cases} 1, & a \geq 0 \\ -1, & a < 0 \end{cases},$$

and if $b = \varepsilon \prod_{i=1}^{k} p_i^{v_p(b)}$ with $p_i \in \mathbb{P}$ and $\varepsilon \in \{\pm 1\}$ we have

$$\left(\frac{a}{b}\right) = \left(\frac{a}{\varepsilon}\right) \prod_{i=1}^{k} \left(\frac{a}{p_i}\right)^{v_p(b)}$$

where $\left(\frac{a}{p}\right)$ is the **Legendre symbol** if p is an odd prime and

$$\left(\frac{a}{2}\right) = \begin{cases} 0, & a \equiv 0 \bmod 2 \\ 1, & a \equiv \pm 1 \bmod 8 \\ -1, & a \equiv \pm 3 \bmod 8 \end{cases}.$$

Since Euclid it is known that there are infinitely many primes. When considering primes in arithmetic progressions, it is clear that there can be at most one prime p with $p \equiv a \bmod m$ if $\gcd(a, m) > 1$. In the other case we have **Dirichlet's theorem on primes in arithmetic progressions**, cf. [Apo76]:

Theorem 0.5.2 (Dirichlet's theorem on primes in arithmetic progressions) *If a and m are coprime, there are infinitely many primes $p \equiv a \bmod m$.*

Although some special cases can be proven in an elementary way (compare [HW08]), known proofs of the general statement require the theory of Dirichlet L-functions (compare Paragraph 0.5.4). In Appendix A.1 we will review the cases in which an elementary proof of some kind is possible.

In some places we will need the Bernoulli numbers.

Definition 0.5.3 The **Bernoulli numbers** B_k are defined via the condition

$$\frac{t}{e^t - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} t^k.$$

More on Bernoulli numbers can be found in [Coh07b]. We call a prime p **regular** if p does not divide the numerators of the even Bernoulli numbers B_k with $k \leq p - 3$. The first irregular prime is 37. For more about regular primes see [Was97]. We will mention an equivalent condition for regular primes in Paragraph 0.5.6.

There are two particular Diophantine equations, that we will try to solve. The first of them is the Pythagorean equation $x^2 + y^2 = z^2$. Solutions (a, b, c) of this equation are called **Pythagorean triples**. A Pythagorean triple (a, b, c) is called **primitive** if $\gcd(a, b, c) = 1$. We can restrict our attention to the case that b is even and a is odd. In this case we have:

Theorem 0.5.4 (characterization of Pythagorean triples) *A triple $(a, b, c) \in \mathbb{N}^3$ is a primitive Pythagorean triple if and only if there are $m, n \in \mathbb{N}$ with $m > n$ and $\gcd(m, n) = 1$ such that*

$$a = m^2 - n^2, \qquad b = 2mn, \qquad c = m^2 + n^2.$$

This can be proved in many different ways, cf. [SF07, Conf]. We will review this equation in Chapters I.7, I.8, and I.12.

The second important equation is Mordell's equation.

Definition 0.5.5 Let $k \in \mathbb{Z}$. The equation $y^2 = x^3 + k$ is called **Mordell's equation**.

We are mostly interested in integer solutions of Mordell's equation. For results about rational solutions see [Mor66]. There are a lot of methods to attack this problem, some of which we will see later. In [Conb], Mordell's equation is examined with some basic methods and a method we will see later. See [Apo76] for another basic result and [Coh07a] for advanced results.

Some values k for which Mordell's equation is not solvable (taken from [Conb, Apo76]) are

$$-73, -65, -57, -37, -24, -17, -6, -5, 6, 7, 11, 23, 35, 45, 46, 87.$$

We will later learn about two methods to examine this equation. One of them will yield a full solution for Mordell's equation for $|k| \leq 10\,000$ (with the help of computer algebra systems).

We need one more lemma about equivalence classes. Since this is not standard, we will give a proof.

Lemma 0.5.6 *Let $\alpha_1, \ldots, \alpha_n$ be representatives of the congruence classes $2, 4, \ldots, 2n$ modulo $(2n + 2)$ and $b \in \mathbb{Z}$ with $\gcd(b, 2n + 2) = 1$. Then $b\alpha_1, \ldots, b\alpha_n$ are representatives of the congruence classes $2, 4, \ldots, 2n$ modulo $(2n + 2)$.*

Proof. It is clear that $b\alpha_i$ and $b\alpha_j$ are incongruent modulo $(2n + 2)$ for all $i \neq j$. Since $b\alpha_i$ is clearly even for all i, we only have to show that $b\alpha_i \not\equiv 0 \bmod (2n + 2)$ for all i. But if $b\alpha_i \equiv 0 \bmod (2n + 2)$ and b is coprime to $2n + 2$, we also have $\alpha_i \equiv 0 \bmod (2n + 2)$, which is a contradiction. q.e.d.

0.5.2 Continued Fractions

Next we consider continued fractions. See [SF07] for proofs and more about continued fractions.

Definition 0.5.7 A (finite) **continued fraction** is a symbol of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}}$$

where $a_0 \in \mathbb{Z}$ and $a_i \in \mathbb{N}$ for $i \geq 1$. We will denote continued fractions by $\langle a_0; a_1, \ldots, a_n \rangle$. A continued fraction is called **normalized** if $a_n \neq 1$.

Continued fractions are yet another way to represent rational numbers. To any given rational number α there are exactly two continued fractions whose value is α, and exactly one of them is normalized.

Continued fractions can be obtained with the **Euclidean Algorithm**: If a and b are nonzero integers, let

$$\begin{aligned} a &= a_0 b + r_1 \\ b &= a_1 r_1 + r_2 \\ r_1 &= a_2 r_2 + r_3 \\ &\vdots \\ r_{n-2} &= a_{n-1} r_{n-1} + r_n \\ r_{n-1} &= a_n r_n \end{aligned}$$

be the Euclidean Algorithm for a and b. Then the normalized continued fraction of $\frac{a}{b}$ is $\langle a_0; a_1, \ldots, a_n \rangle$.

Completely analogous (and thus omitted here) is the definition of infinite continued fractions. These fractions represent irrational numbers. We will write infinite continued fractions in the form $\langle a_0; a_1, a_2, \ldots \rangle$. In both the finite and the infinite case the integers a_i are called **partial quotients**.

It is easy to compute the continued fraction expansion (or at least a finite number of partial quotients) of an irrational number y by repeatedly writing

$$y = \lfloor y \rfloor + \{y\} = \lfloor y \rfloor + \cfrac{1}{\left\lfloor \frac{1}{\{y\}} \right\rfloor + \left\{ \frac{1}{\{y\}} \right\}} = \cdots .$$

Here and in the following, $\lfloor x \rfloor$ denotes the **floor function**, i.e., the largest integer $a \leq x$. We denote by $\{x\} = x - \lfloor x \rfloor$ the **fractional part** of x. Since this notation might become unclear when also dealing with sets, we will explicitely mention this notation everytime we use it.

For example, the continued fraction of e is given by

$$\langle 1; 0, 1, 1, 2, 1, 1, 4, 1, 1, 6, \ldots, 1, 1, 2n, \ldots \rangle.$$

See [Coh06] for a nice proof of this fact.

For the opposite direction, i.e., if one wants to get the real number that is represented by the continued fraction $\alpha = \langle a_0; a_1, a_2, \ldots \rangle$, one can use the recurrence formula

$$p_{-1} := 1, p_0 := a_0, p_n := a_n p_{n-1} + p_{n-2}, \quad q_{-1} := 0, q_0 := 1, q_n := a_n q_{n-1} + q_{n-2}.$$

Then for any $n \in \mathbb{N}$ we have $\langle a_0; a_1, \ldots, a_n \rangle = \frac{p_n}{q_n}$ and $\lim_{n\to\infty} \frac{p_n}{q_n} = \alpha$. The fraction $\frac{p_n}{q_n}$ is called the ***n*-th convergent** of α.

Similarly to the case of decimal fractions there is a nice characterization of periodic continued fractions due to Euler and Lagrange:

Theorem 0.5.8 *Let α be irrational. The continued fraction of α is **periodic** (i.e., of the form $\langle a_0; a_1, \ldots, a_m, \overline{b_1, \ldots, b_n} \rangle$) with $n \geq 1$ if and only if α is **algebraic** of degree 2, i.e., $\alpha = r + q\sqrt{d}$ with $r, q \in \mathbb{Q}, d \in \mathbb{Z} \backslash \{1\}$ squarefree.*

0.5.3 Additive Number Theory

Additive number theory is concerned with questions involving **sumsets**:

Definition 0.5.9 Let A, B be subsets of an additive structure (mostly $\mathbb{N}$ or $\mathbb{Z}/m\mathbb{Z}$ for some m). The sumset of A and B is

$$A + B := \{a + b : a \in A, b \in B\}.$$

There are two possible types of problems in this field: Determining sumsets (or their properties) for given sets or deducing information about the sets A and B when a sumset is given.

The classical problem of the first variant is the following: Given a set $S \subset \mathbb{N}_0$ and a natural number k, determine which natural numbers n can be written as

a sum of k elements of S (and determine, if possible, the number of representations). Questions of these type are dealt with in [Nat96b]. For the inverse question, i.e., deducing information about A and B when the sumset of A and B is given, see [Nat96a]. We will see both kinds of problems. Here we just mention some classical problems and results.

Three of the most classical results are the two-squares theorem, the three-squares theorem, and the four-squares theorem. Proofs and more about these results can be found in [Nat96b, SF07] and, for the case of the three-squares theorem, in [For08]. Note that in all three cases the squares involved may be squares of integers, i.e., the square 0 is allowed.

Theorem 0.5.10 (two-squares theorem) *A natural number $n \in \mathbb{N}$ can be written as a sum of two squares if and only if $2|v_p(n)$ for every prime $p \equiv 3 \bmod 4$.*

Theorem 0.5.11 (three-squares theorem) *A natural number $n \in \mathbb{N}$ can be written as a sum of three squares if and only if n can be written in the form $n = 4^a m$ with $4 \nmid m$ and $m \not\equiv 7 \bmod 8$.*

Theorem 0.5.12 (four-squares theorem) *Every natural number can be written as a sum of four squares.*

In general, for given k, **Waring's problem** is concerned with the determination of the smallest natural number $g(k)$ such that any natural number can be written as a sum of $g(k)$ k-th powers, cf. [Nat96b, Vau97]

Problem 0.5.13 (Waring's problem) *Given a natural number k, what is the smallest natural number $g(k)$ such that any natural number can be written as a sum of $g(k)$ k-th powers?*

The existence of $g(k)$ has been shown by Hilbert [Hil09]. The value $g(k)$ is known for all k, cf. [Niv44]. The first ten values are

$$1, 4, 9, 19, 37, 73, 279, 548, 1079, 2132.$$

A variant of Waring's problem asks for the smallest natural number $G(k)$ such that all sufficiently large natural numbers can be written as a sum of $G(k)$ k-th powers. This problem is unsolved for most k, until now only the values $G(1) = 1$, $G(2) = 4$, and $G(4) = 16$ are known. For other values of k only bounds are known, cf. [Vau97]. This variant of Waring's problem can be attacked with a method we will present in Chapter I.10.

Another problem is concerned with the sumset of primes. These are Goldbachs problems, cf. [Nat96b, Vau97].

Problem 0.5.14 (ternary Goldbach problem) *Can every odd number greater than* 5 *be written as a sum of three primes?*

Problem 0.5.15 (binary Goldbach problem) *Can every even number greater than* 2 *be written as a sum of two primes?*

While the ternary Goldbach problem can be attacked with the method mentioned in Chapter I.10 (and is indeed, as it seems, solved, cf. [Hel13, Hel12]), the binary Goldbach problem is still unsolved.

An example of a result about sumsets in finite fields is the **Cauchy-Davenport theorem**, see [Nat96a]:

Theorem 0.5.16 (Cauchy-Davenport theorem) *Let* $p \in \mathbb{P}$ *and* $A, B \subset \mathbb{Z}/p\mathbb{Z}$ *be nonempty. Then*

$$|A+B| \geq \min\{p, |A|+|B|-1\}.$$

We will see generalizations and inverse results of the Cauchy-Davenport theorem in Chapter I.6.

0.5.4 The Riemann ζ-function and Dirichlet series

One of the most important functions in number theory is the Riemann ζ-function. As usual we will denote the complex argument by s.

Definition 0.5.17 The **Riemann ζ-function** is defined via the series

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s}.$$

This series converges for $\Re(s) > 1$. We state some results about $\zeta(s)$ and recall the Riemann hypothesis since both the results (as well as generalisations) and the conjecture will recur later. For proofs see [Brü95].

Theorem 0.5.18 *The Riemann ζ-function has the following properties:*

- *For $\Re(s) > 1$ the Riemann ζ-function admits an **Euler product***

$$\zeta(s) = \prod_{p \in \mathbb{P}} \frac{1}{1-p^{-s}}.$$

- *The following functional equation holds:*

$$\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \pi^{-\frac{1-s}{2}}\Gamma\left(\frac{1-s}{2}\right)\zeta(1-s).$$

Here Γ *is the* ***Γ-function***

$$\Gamma(s) := \int_0^\infty e^{-t} t^{s-1} \, \mathrm{d}t.$$

The functional equation gives a meromorphic continuation of $\zeta(s)$ *to* $\mathbb{C}$*. More precisely,* $\zeta(s)$ *is holomorphic in* $\mathbb{C}\backslash\{1\}$ *and has a simple pole with residue* 1 *at* $s = 1$*.*

- *The Riemann* ζ*-function has zeros at* $-2n$ *for all* $n \in \mathbb{N}$*. These are called* ***trivial zeros****.*

One of the most important open problems in mathematics is the **Riemann hypothesis**:

Conjecture 0.5.19 (Riemann hypothesis) *All nontrivial zeros of* $\zeta(s)$ *have real part* $\frac{1}{2}$*.*

For an introduction to the Riemann hypothesis see [Bom06]. The Riemann hypothesis is one of the Millennium problems, cf. [CJW06]. For some equivalent formulations see [CF, Ban, BCRW08].

More about the Riemann ζ-function and the Riemann hypothesis in particular can be found in [Tit86, BCRW08].

We take a brief look at a generalization of the Riemann ζ-function. Let $(a_n)_{n\in\mathbb{N}}$ be a complex sequence. The **Dirichlet series** associated to this sequence is the series $\sum_{n=1}^\infty \frac{a_n}{n^s}$ with $s \in \mathbb{C}$. Properties of Dirichlet series (for example regarding its convergence) can be found in [Brü95].

A special case are Dirichlet L-functions. Recall that a **Dirichlet character modulo** m is a function $\chi : \mathbb{Z} \to \mathbb{C}$ such that

- $\chi(n) = \chi(m + n)$ for all $n \in \mathbb{Z}$,
- $\chi(ab) = \chi(a)\chi(b)$ for all $a, b \in \mathbb{Z}$,
- $\chi(n) = 0$ if and only if $\gcd(m, n) > 1$.

A **Dirichlet *L*-function** is a Dirichlet series associated to a Dirichlet character, i.e., a series of the form $\sum_{n=1}^\infty \frac{\chi(n)}{n^s}$.

There is an analogue of the Riemann hypothesis for Dirichlet L-functions, the **generalized Riemann hypothesis**. This asserts that for any complex s with $L(s, \chi) = 0$ and $\Re(s) \in [0, 1]$ we have in fact $\Re(s) = \frac{1}{2}$.

0.5.5 Transcendental Number Theory

We will need a few results of transcendental number theory. All proofs of the results in this paragraph can be found in [Bak90].

One of the starting points of transcendental number theory is Liouville's theorem about approximation of algebraic numbers.

Theorem 0.5.20 (Liouville's theorem) *Let α be an algebraic number of degree $d > 1$. Then there is a number $c > 0$ depending on α such that*

$$\left|\alpha - \frac{a}{b}\right| > \frac{c}{b^d}$$

holds for all rational numbers $\frac{a}{b}$ with $b > 0$.

Liouville's theorem can be used to show that the number $\sum_{k=1}^{\infty} 10^{-k!}$ is transcendental. The lower bound in Liouville's theorem means that an algebraic number cannot be approximated too well. The converse is not true. For example, Mahler [Mah53] showed that

$$\left|\pi - \frac{a}{b}\right| > b^{-42}$$

for all $a, b \in \mathbb{N}$. The exponent 42 has later been reduced, the best known bound is approximately 7.60630852, see [Sal08]. However, it is possible to get lower bounds for approximations of some transcendental numbers using results about linear forms in logarithms. To state this, recall that the **height** of an algebraic number α with minimal polynomial $f = a_n x^n + \cdots + a_0$ is defined as the number $\max\{|a_i| : 0 \leq i \leq n\}$.

Theorem 0.5.21 (Baker's theorem on linear forms in logarithms) *Let $\alpha_1, \ldots, \alpha_n$ be nonzero algebraic numbers with degrees at most d and heights at most h. Let further $\beta_0, \ldots, \beta_n$ be algebraic numbers with degrees at most d and heights at most H with $H \geq 2$. Define*

$$L := \beta_0 + \beta_1 \log(\alpha_1) + \cdots + \beta_n \log(\alpha_n).$$

If $L \neq 0$ we have $|L| > H^{-C}$ where C is a constant depending on n, d, and h.

With Baker's theorem one can get lower bounds on approximations of transcendental numbers that can be written as the quotient of two logarithms. We will see an example in Section II.4.4.

Although the set of algebraic numbers is countable and the set of transcendental numbers is uncountable, it is in general hard to give examples of transcendental numbers. The **Lindemann-Weierstraß theorem** can be used to prove that certain numbers are transcendental.

Theorem 0.5.22 (Lindemann-Weierstraß theorem) *Let $\alpha_1, \ldots, \alpha_n$ be distinct algebraic numbers and let $\beta_1, \ldots, \beta_n$ be nonzero algebraic numbers. Then*

$$\beta_1 e^{\alpha_1} + \cdots + \beta_n e^{\alpha_n} \neq 0.$$

With the Lindemann-Weierstraß theorem one can show that $e^\alpha, \sin(\alpha), \cos(\alpha)$ and $\tan(\alpha)$ are transcendental for any algebraic $\alpha \neq 0$, and that $\log(\alpha)$ is transcendental for any algebraic $\alpha \in \mathbb{C}\backslash\mathbb{R}_{\leq 0}, \alpha \neq 1$ (here log denotes the principal value of the complex logarithm).

0.5.6 Number Fields

To conclude the basics in number theory we will recall some concepts and results of algebraic number theory. For the understanding of this thesis we will just need a few basic definitions and some results. Therefore we will just mention enough to understand the results. All additional information will be mentioned without explaining the concepts. Proofs of all results and more on algebraic number theory can (unless stated otherwise) be found in [Neu92].

We will start with number fields here and will deal with Galois theory and valuations in the next paragraphs.

Definition 0.5.23 A **number field** is a finite extension of $\mathbb{Q}$.

We are interested in the ring of integers of a number field K.

Definition 0.5.24 Let S/R be a ring extension. An element $s \in S$ is called **integral** over R if there is a monic polynomial $f \in R[x]$ such that $f(s) = 0$. The set $\{s \in S : s \text{ is integral over } R\}$ is called **integral closure** of R in S. If $\mathrm{Quot}(R)$ denotes the quotient field of R and the integral closure of R in $\mathrm{Quot}(R)$ is R itself, then R is called **integrally closed**.

If K is a number field, the integral closure of $\mathbb{Z}$ in K, denoted by $\mathcal{O}_K$, is called the **ring of integers** of K. An important property of rings of integers is the following: If $\alpha \in \mathcal{O}_K$ is rational, then in fact $\alpha \in \mathbb{Z}$.

The ring $\mathcal{O}_K$ is equipped with a canonical norm function. For this we consider the embeddings of K in $\mathbb{C}$. There are exactly $n = [K : \mathbb{Q}]$ different embeddings,

where $[K : \mathbb{Q}]$ denotes the degree of the field extension $K/\mathbb{Q}$. We call an embedding **real** if its image is contained in $\mathbb{R}$. Otherwise the embedding is called **complex**. Since embeddings correspond to roots of the minimal polynomial of a primitive element of K, the complex embeddings come in pairs. We will denote the number of real embeddings by r and the number of complex embeddings by $2s$ and call complex embeddings corresponding to conjugate roots **conjugate**. K is called **totally real** if $s = 0$, i.e., if all embeddings are real. If $\sigma_i, i = 1, \ldots, n$, are the embeddings of K in $\mathbb{C}$, the **norm** of an element $\beta \in \mathcal{O}_K$ is

$$N_{\mathbb{Q}}^K(\beta) := \prod_{i=1}^{n} \sigma_i(\beta).$$

The ring of integers $\mathcal{O}_K$ is in fact a free $\mathbb{Z}$-module of rank n (i.e., $\mathcal{O}_K$ has a basis and each basis of $\mathcal{O}_K$ has exactly n elements). If $\{b_1, \ldots, b_n\}$ is such a basis, the **discriminant** Δ_K of K is defined to be

$$\Delta_K = \det \begin{pmatrix} \sigma_1(b_1) & \cdots & \sigma_1(b_n) \\ \vdots & \ddots & \vdots \\ \sigma_n(b_1) & \cdots & \sigma_n(b_n) \end{pmatrix}^2.$$

Recall that a number $\zeta \in \mathbb{C}$ is called ***n*-th root of unity** if $\zeta^n = 1$. If n is the smallest exponent with that property, ζ is called **primitive *n*-th root of unity**. For all n, the number $\zeta_n = e^{\frac{2\pi i}{n}}$ is a primitive n-th root of unity.

Example 0.5.25 Let $\alpha = \sqrt[3]{2}$ and $K = \mathbb{Q}(\alpha)$. Then $[K : \mathbb{Q}] = 3$ and the minimal polynomial of α is $f(x) = x^3 - 2$. The roots of f are $\alpha, \zeta\alpha$, and $\zeta^2\alpha$ where ζ is a primitive third root of unity, e.g., $\zeta = \frac{-1+\sqrt{-3}}{2}$. Hence the discriminant of f is

$$\Delta_f = \prod_{i<j} (\alpha\zeta^i - \alpha\zeta^j)^2 = -108.$$

The three embeddings of K in $\mathbb{C}$ are given by $\sigma_1(\alpha) = \alpha, \sigma_2(\alpha) = \zeta\alpha$ and $\sigma_3(\alpha) = \zeta^2\alpha$. The embedding σ_1 is real while σ_2 and σ_3 are complex, i.e., we have $r = s = 1$. A basis of $\mathcal{O}_K$ is given by $\{1, \alpha, \alpha^2\}$ and thus

$$\Delta_K = \det \begin{pmatrix} 1 & \alpha & \alpha^2 \\ 1 & \zeta\alpha & \zeta^2\alpha^2 \\ 1 & \zeta^2\alpha & \zeta\alpha^2 \end{pmatrix}^2 = -108 = \Delta_f.$$

If $\beta \in K$ is of the form $\beta = a + b\alpha + c\alpha^2$, then $N_{\mathbb{Q}}^K(\beta) = a^3 + 2b^3 + 4c^3 - 6abc$.

The equality of the discriminants in the above example was no coincidence. If a basis of $\mathcal{O}_K$ is given by $\{1, \alpha, \ldots, \alpha^{n-1}\}$ and f is the minimal polynomial of α, then we always have $\Delta_K = \Delta_f$.

We are in particular interested in (the structure of) units and ideals of $\mathcal{O}_K$. Further, if K has degree 2, we want to know what additional structure these rings exhibit. We start with the investigation of the case $[K : \mathbb{Q}] = 2$. The following theorem gives the structure of $\mathcal{O}_K$ and the discriminant Δ_K. Recall that in this case the number field K is given by $K = \mathbb{Q}(\sqrt{d}) = \{x + y\sqrt{d} : x, y \in \mathbb{Q}\}$ for some $d \in \mathbb{Z}\backslash\{0, 1\}$ that is not a square. Since $\mathbb{Q}(\sqrt{d}) = \mathbb{Q}(\sqrt{d'})$ if $d = d' \cdot r^2$ with $r \in \mathbb{Z}$, we can further assume that d is squarefree.

Theorem 0.5.26 *Let $K = \mathbb{Q}(\sqrt{d})$ with squarefree d and $d \neq 0, 1$.*

- *The ring of integers of K is*

$$\mathcal{O}_K = \begin{cases} \mathbb{Z}[\frac{1+\sqrt{d}}{2}] := \left\{ \frac{a+b\sqrt{d}}{2} : a, b, \in \mathbb{Z}, a \equiv b \bmod 2 \right\}, & \text{if } d \equiv 1 \bmod 4 \\ \mathbb{Z}[\sqrt{d}] := \left\{ a + b\sqrt{d} : a, b, \in \mathbb{Z} \right\}, & \text{if } d \equiv 2, 3 \bmod 4 \end{cases}.$$

- *$\mathcal{O}_K$ is Euclidean if and only if*

$$d \in \{-1, -2, -3, -7, -11, 2, 3, 5, 6, 7, 11, 13, 17, 19, 21, 29, 33, 37, 41, 57, 73\}.$$

- *For negative d, $\mathcal{O}_K$ is a unique factorization domain if and only if*

$$d \in \{-1, -2, -3, -7, -11, -19, -43, -67, -163\}.$$

- *The ring $\mathcal{O}_K$ contains*
  - *exactly the sixth roots of unity if and only if $d = -3$,*
  - *exactly the fourth roots of unity if and only if $d = -1$,*
  - *exactly the second roots of unity if and only if $d \neq -1, -3$.*

- *The discriminant of K is*

$$\Delta_K = \begin{cases} d, & \text{if } d \equiv 1 \bmod 1 \\ 4d, & \text{if } d \equiv 2, 3 \bmod 4 \end{cases}.$$

For the first, fourth and fifth statement of the theorem see [Sch07a]. For the second statement see [Leu96, Hal97] and the references mentioned there. The third statement is proved in [Sta67].

In the following, whenever we work with $\mathbb{Q}(\sqrt{d})$ or $\mathbb{Z}[\sqrt{d}]$ we assume that $d \neq 0, 1$ and d is not a square. Since we sometimes need to work with non-squarefree d we will not exclude this case but due to the comments above, in most results we will further assume that d is squarefree.

The ring of integers of a number field K has some further important properties.

Definition 0.5.27 Let R be an integral domain. R is called **Dedekind domain** if the following three properties hold:

- R is **Noetherian**, i.e., every ideal of R is finitely generated.
- R is integrally closed.
- Every nonzero prime ideal is a maximal ideal.

There are many other characterizations and properties of Dedekind domains (see [Ash03]), in particular a Dedekind domain is a principal ideal domain if and only if it is a unique factorization domain. One can show that for any number field K the ring of integers $\mathcal{O}_K$ is a Dedekind domain.

Now we examine the structure of the unit group $\mathcal{O}_K^*$. First we note that an element $\alpha \in \mathcal{O}_K$ is a unit if and only if $N_{\mathbb{Q}}^K(\alpha) = \pm 1$. In the case $\mathcal{O}_K = \mathbb{Z}[\sqrt{d}]$ this means that $\alpha = x + y\sqrt{d}$ is a unit if and only if $x^2 - dy^2 = \pm 1$. If d is negative, then the units in $\mathbb{Z}[\sqrt{d}]$ are exactly the roots of unity mentioned in Theorem 0.5.26, see [Sch07a]. For general number fields K the unit group can be infinite.

Definition 0.5.28 Let G be a finitely generated group. The **rank** of G is the cardinality of the smallest subset $S \subset G$ that generates G.

Theorem 0.5.29 (Dirichlet's unit theorem) *The unit group of $\mathcal{O}_K$ is the direct product of the roots of unity in K (this is a finite, cyclic group) and an abelian group of rank $r + s - 1$.*

Let us now take a look at the structure of the ideals. Here by structure we mean that we want to know how far away the ring $\mathcal{O}_K$ is from being a principal ideal domain. For this, we first need to turn the set of ideals of $\mathcal{O}_K$ into a group, i.e., we need inverses of ideals.

Definition 0.5.30 A **fractional ideal** $\mathfrak{a}$ of $\mathcal{O}_K$ is a finitely generated $\mathcal{O}_K$-submodule of K, i.e., a finitely generated subgroup of K such that $ar \in \mathfrak{a}$ for all $a \in \mathfrak{a}, r \in \mathcal{O}_K$. A fractional ideal is called **principal** if it is generated by a single element.

Since ideals in $\mathcal{O}_K$ are finitely generated, ideals are fractional ideals. Analogously to usuual ideals, the product of two fractional ideals $\mathfrak{a}$ and $\mathfrak{b}$ is defined as

$$\mathfrak{a} \cdot \mathfrak{b} = \{a_1 b_1 + \cdots + a_n b_n : a_i \in \mathfrak{a}, b_i \in \mathfrak{b}, n \in \mathbb{N}\}.$$

The set of all nonzero fractional ideals forms a group (together with the above multiplication), see [Ash03]. We denote this group by I_K. Let P_K be the subgroup of principal ideals. The **ideal class group** $\mathcal{C}l_K$ is defined as $\mathcal{C}l_K := I_K / P_K$. We are interested in the cardinality of $\mathcal{C}l_K$. The number $h_K := |\mathcal{C}l_K|$ is called the **class number** of K. If $h_K = 1$, the ring $\mathcal{O}_K$ is a principal ideal domain, and the smaller h_K is, the closer $\mathcal{O}_K$ is to being a principal ideal domain. Regarding the finiteness of h_K we have the following result:

Theorem 0.5.31 (finiteness of the class number) *For any number field K we have $h_K < \infty$.*

Class numbers yield another characterization of regular primes: A prime p is regular if and only if p does not divide $h_{\mathbb{Q}(\zeta_p)}$ (here ζ_p is a primitive p-th root of unity), see [Edw77].

One of the biggest benefits of Dedekind domains is the existence of a prime ideal decomposition. In the case of rings of integers we have:

Theorem 0.5.32 *Let K be a number field and $\mathcal{O}_K$ its ring of integers. Every nontrivial ideal $\mathfrak{a} \neq (0), (1)$ of $\mathcal{O}_K$ has a unique decomposition (up to order)*

$$\mathfrak{a} = \mathfrak{p}_1^{v_1} \cdots \mathfrak{p}_r^{v_r}$$

where $\mathfrak{p}_i$ are prime ideals of $\mathcal{O}_K$ and $v_i \in \mathbb{N}$.

Let $p \in \mathbb{Z}$ be a prime and (p) the principal ideal in the ring of integers of some number field K, i.e. $(p) = \{ap : a \in \mathcal{O}_K\}$. Let $(p) = \mathfrak{p}_1^{v_1} \cdots \mathfrak{p}_r^{v_r}$ be the decomposition of (p) in prime ideals of $\mathcal{O}_K$ as in Theorem 0.5.32. Then p is called

- **inert** if $r = 1$ and $v_1 = 1$, i.e., (p) is a prime ideal in $\mathcal{O}_K$,
- **split** if $r > 1$ and $v_i = 1$ for all i,
- **ramified** if $v_i > 1$ for some i.

It is known that a prime p ramifies in $\mathcal{O}_K$ if and only if p divides Δ_K. For quadratic number fields K, the splitting behaviour of a prime p in $\mathcal{O}_K$ can be determined with the Kronecker symbol, cf. [Sch07a].

Theorem 0.5.33 *Let $d \in \mathbb{Z}\backslash\{1\}$ be squarefree, $K = \mathbb{Q}(\sqrt{d})$ and $p \in \mathbb{P}$. Then we have the following:*

- *p is inert in $\mathcal{O}_K$ if and only if $\left(\frac{\Delta_K}{p}\right) = -1$.*
- *p is split in $\mathcal{O}_K$ if and only if $\left(\frac{\Delta_K}{p}\right) = 1$.*
- *p is ramified in $\mathcal{O}_K$ if and only if $\left(\frac{\Delta_K}{p}\right) = 0$.*

For the ring $\mathbb{Z}[i]$ of Gauß integers, we thus get:

Theorem 0.5.34 *Let $p \in \mathbb{P}$ be a prime.*

- *p is inert (i.e., a prime in $\mathbb{Z}[i]$) if and only if $p \equiv 3 \bmod 4$.*
- *p is split if and only if $p \equiv 1 \bmod 4$, in this case p can be decomposed in two non-associated primes.*
- *p is ramified if and only if $p = 2$ (more precisely we have $2 = (1-i)(1+i)$ and $1-i, 1+i$ are primes that are associated).*

We conclude this paragraph with the definition of the **Dedekind ζ-function**, a generalization of the Riemann ζ-function to number fields.

Definition 0.5.35 Let K be a number field. The Dedekind ζ-function of K is defined as

$$\zeta_K(s) = \sum_{\mathfrak{a} \lhd \mathcal{O}_K} \frac{1}{N(\mathfrak{a})^s}$$

where $\mathfrak{a}$ runs over all ideals of $\mathcal{O}_K$ and $N(\mathfrak{a}) = |\mathcal{O}_K/\mathfrak{a}|$ denotes the **norm** of the ideal $\mathfrak{a}$.

Note that $\zeta_{\mathbb{Q}}(s) = \zeta(s)$ is the Riemann ζ-function. The Dedekind ζ-function has properties similar to the Riemann ζ-function such as meromorphic continuation and functional equation, see [Neu92].

0.5.7 Galois theory

We need some results from Galois theory. Since we need this only in a special case, we will skip the general theory and will just cover (special) number fields here. For general results and proofs see [Bos05].

Definition 0.5.36 Let K be a number field. The extension $K/\mathbb{Q}$ is called **Galois extension** if the minimal polynomials (over $\mathbb{Q}$) of all elements $\alpha \in K$ factor into linear factors in K.

If K is a number field, a $\mathbb{Q}$**-automorphism** of K is an automorphism $\sigma : K \to K$ such that $\sigma(\mathbb{Q}) = \mathbb{Q}$. The group of $\mathbb{Q}$-automorphisms of a Galois extension K, the so-called **Galois group**, is denoted by $\mathrm{Gal}(K/\mathbb{Q})$.

Example 0.5.37 Let $\zeta_n = e^{\frac{2\pi i}{n}}$ be a primitive n-th root of unity and $K_n = \mathbb{Q}(\zeta_n)$. Then $K_n/\mathbb{Q}$ is a Galois extension and $[K_n : \mathbb{Q}] = \varphi(n)$. The $\mathbb{Q}$-automorphisms σ_j of K_n are given by

$$\sigma_j(\zeta_n) = \zeta_n^j, \quad 1 \leq j \leq n, \gcd(j, n) = 1,$$

i.e., $\mathrm{Gal}(K/\mathbb{Q}) \cong (\mathbb{Z}/n\mathbb{Z})^*$.

If K is a number field and $G \subset \mathrm{Gal}(K/\mathbb{Q})$, the field

$$K^G := \{\alpha \in K : \sigma(\alpha) = \alpha \text{ for all } \sigma \in G\}$$

is called the **fixed field** of G in K. For Galois extensions there is an important result about fixed fields:

Theorem 0.5.38 *Let $K/\mathbb{Q}$ be a Galois extension. Then $K^{\mathrm{Gal}(K/\mathbb{Q})} = \mathbb{Q}$.*

0.5.8 Valuations

In Chapter I.1 we will deal with p-adic numbers and p-adic valuations. These are just special cases of valuations and absolute values which we will use again in Chapter I.4. Here we mention the concepts and give some basic results, for deeper results and proofs see [Neu92].

Definition 0.5.39 Let K be a field.

- An **absolute value** is a function $|\cdot| : K \to \mathbb{R}_{\geq 0}$ such that for $x, y \in K$ the following properties hold:
  - $|x| = 0 \Leftrightarrow x = 0$
  - $|xy| = |x|\,|y|$
  - $|x + y| \leq |x| + |y|$
- A **valuation** is a function $v : K \to \mathbb{R} \cup \infty$ such that for $x, y \in K$ the following properties hold:
  - $v(x) = \infty \Leftrightarrow x = 0$
  - $v(xy) = v(x) + v(y)$
  - $v(x + y) \geq \min\{v(x), v(y)\}$

An absolute value $|\cdot|$ is called **trivial** if $|x| = 1$ for all $x \in K\backslash\{0\}$. A valuation v is called **trivial** if $v(a) = 0$ for all $x \in K\backslash\{0\}$. We will always assume that absolute values and valuations are nontrivial.

An absolute value $|\cdot|$ on a field K is called **non-Archimedean** if for any $x, y \in K$ the **strong triangle inequality** $|x+y| \leq \max\{|x|, |y|\}$ holds. If this is not the case, the absolute value is called **Archimedean**.

There is a 1-to-1 correspondence between valuations and non-Archimedean absolute values: If $v(x)$ is a valuation and $q > 1$, then $|x| = q^{-v(x)}$ is a non-Archimedean absolute value. Conversely, $v(x) = -\log_q(|x|)$ is a valuation if $|x|$ is a non-Archimedean absolute value and $q > 1$.

Absolute values (and valuations) induce a metric d on K via $d(x,y) = |x-y|$. Hence each absolute value induces a topology on K. We call two absolute values $|\cdot|_1$ and $|\cdot|_2$ **equivalent** if they induce the same topology. This is equivalent to the existence of a real number $s > 0$ with $|x|_1 = |x|_2^s$ (respectively $v_1(x) = sv_2(x)$ for the corresponding valuations) for all $x \in K$.

Ostrowski's theorem characterizes the equivalence classes of absolute values on $\mathbb{Q}$. This can be generalized to number fields, see [Cond]:

Theorem 0.5.40 (Ostrowski's theorem) *Let $|\cdot|$ be a nontrivial absolute value on a number field K. Then $|\cdot|$ is equivalent to exactly one of the following absolute values:*

- *An absolute value defined by a real embedding of K in $\mathbb{C}$.*
- *An absolute value defined by a pair of complex embeddings of K in $\mathbb{C}$.*
- *An absolute value defined by a prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$.*

Here the absolute value of an embedding σ is defined by $|x|_\sigma := |\sigma(x)|$ where $|\cdot|$ is the standard absolute value on $\mathbb{R}$ or $\mathbb{C}$, respectively. The absolute values defined through prime ideals are completely analogous to the p-adic absolute value that we will see in Chapter I.1. Since we do not need to know how these valuations exactly look like we omit a precise definition.

If $|\cdot|$ is an absolute value on K, then $(K, |\cdot|)$ (or simply K) is called a **valued field**. We wish to obtain a field extension $K_{|\cdot|}$ of K such that all Cauchy sequences (with respect to $|\cdot|$) converge to an element in $K_{|\cdot|}$. This process is completely analogous to the construction of $\mathbb{R}$. Moreover, we will see this construction again in Chapter I.1. Since we do not need the basics about valuations in Chapter I.1 but in Chapter I.4 (and despite the fact that the chapters in Part 1 have interconnections, we want the basic concepts of each chapter to be understandable without knowledge of previous chapters), we will explain the construction in its generality here and in the special case again in Chapter I.1.

Definition 0.5.41 Let $(K, |\cdot|)$ be a valued field.

- A **Cauchy sequence** in K is a sequence $(a_n)_{n\in\mathbb{N}}$ such that for any $\varepsilon > 0$ there is a $n_0 \in \mathbb{N}$ with $|a_n - a_m| < \varepsilon$ for all $n, m \geq n_0$. We denote the ring of Cauchy sequences by $\mathbf{c}_{|\cdot|}(K)$ and its maximal ideal of sequences that converge to 0 with $\mathbf{0}_{|\cdot|}(K)$.
- The **completion** $\widehat{K}$ of $(K, |\cdot|)$ is defined as the quotient $\mathbf{c}_{|\cdot|}(K)/\mathbf{0}_{|\cdot|}(K)$. If one emphasizes the completion with respect to a valuation v or an absolute value $|\cdot|$, this is denoted by K_v or $K_{|\cdot|}$.
- K is called **complete** if $\widehat{K} = K$.

Lastly, we extend absolute values to specific extensions of K.

Definition 0.5.42 Let $(K, |\cdot|_K)$ be a valued field and L/K a field extension. An absolute value $|\cdot|_L$ on L is called **extension** of $|\cdot|_K$ if the restriction of $|\cdot|_L$ to K is exactly $|\cdot|_K$.

The absolute value $|\cdot|$ has a canonical extension on $\widehat{K}$ via $|x| = \lim_{n\to\infty} |x_n|$ if $x_n \to x$. This extends uniquely to the algebraic closure of $\widehat{K}$ (we will denote the **algebraic closure** of a field K by $\overline{K}$). In general, the algebraic closure of $\widehat{K}$ will not be a complete field. But the completion of the algebraic closure of $\widehat{K}$ is a complete, algebraic closed field. For Archimedean valuations this follows by Ostrowski's theorem (Theorem 0.5.40), for the non-Archimedean case see [Cla10].

0.6 Applied Mathematics and Computer Science

In this section we will cover basics of probability theory and formal languages and we will review the Netwon algorithm.

0.6.1 Measure spaces and Dynamical Systems

Let us recall some definitions from measure theory. For more information see [Kle08].

Definition 0.6.1 Given a set M, a **σ-algebra** $\mathscr{A}$ of M is a subset $\mathscr{A} \subset \mathcal{P}(M)$ such that

- $M \in \mathscr{A}$,
- if $A \in \mathscr{A}$, then $M\backslash A \in \mathscr{A}$,
- if $A_i \in \mathscr{A}$ for $i \in \mathbb{N}$, then $\bigcup_{i\in\mathbb{N}} A_i \in \mathscr{A}$.

A pair $(M, \mathscr{A})$, where $\mathscr{A}$ is a σ-algebra of M, is called a **measurable space** and elements of $\mathscr{A}$ are called **measurable sets**.

Let $(M_1, \mathscr{A}_1)$ and $(M_2, \mathscr{A}_2)$ be two measurable spaces and $f : M_1 \to M_2$ a function. Then f is called **measurable** if for any $A \in \mathscr{A}_2$ we have $f^{-1}(A) \in \mathscr{A}_1$, i.e., preimages of measurable sets are measurable. For a topological space (M, τ) there is a canonical σ-algebra $\mathscr{B}$ on M, the **Borel σ-algebra**. This is generated by all open sets. In this setting, continuous functions are measurable.

Let $(M, \mathscr{A})$ be a measurable space. A function $\mu : \mathscr{A} \to \mathbb{R} \cup \{\infty\}$ is called a **measure** if

- for all $A \in \mathscr{A}$ we have $\mu(A) \geq 0$,
- $\mu(\emptyset) = 0$,
- $\mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$ for pairwise disjoint $A_i \in \mathscr{A}$.

If μ is a measure on $(M, \mathscr{A})$, then $(M, \mathscr{A}, \mu)$ is called a **measure space**. $(M, \mathscr{A}, \mu)$ is called a **probability space** if further $\mu(M) = 1$.

Let $\mathscr{B}$ be the Borel σ-algebra on $\mathbb{R}^n$ and $a = (a_1, \ldots, a_n), b = (b_1, \ldots, b_n) \in \mathbb{R}^n$. Let (a, b) denote the set

$$\{x = (x_1, \ldots, x_n) \in \mathbb{R}^n : x_i \in (a_i, b_i) \text{ for all } i\}.$$

There is a unique measure λ_n on the measurable space $(\mathbb{R}^n, \mathscr{B})$ such that

$$\lambda_n((a, b)) = \prod_{i=1}^{n} (b_i - a_i)$$

if $a_i < b_i$ for all i. This measure is called the **Lebesgue measure**.

If $(M, \mathscr{A}, \mu)$ is a measure space, a set $N \in \mathscr{A}$ with $\mu(N) = 0$ is called **null set**. If a property E is satisfied for all $m \in M \backslash N$ for some null set N, then we say that E is satisfied **almost everywhere** or for **almost all** m.

We are interested in measurable functions between measure spaces that preserve the measure.

Definition 0.6.2 Let $(M, \mathscr{A}, \mu)$ be a measure space. Then a measurable function $T : M \to M$ is called **measure preserving** if $\mu(T^{-1}(A)) = \mu(A)$ for all $A \in \mathscr{A}$. In this case we call μ a ***T*-invariant measure** and $(M, \mathscr{A}, \mu, T)$ a **dynamical system**. If further $(M, \mathscr{A}, \mu)$ is a probability space, we say that $(M, \mathscr{A}, \mu, T)$ is a **dynamical system over a probability space**.

Example 0.6.3 Consider the set $M = [3,5]$ and let $\mathscr{B}$ be the Borel σ-algebra on M (with respect to the usual topology). Let $\mu = 7\lambda$ where λ is the usual Lebesgue measure, i.e., $\mu(M) = 14$. Let T be the map

$$T : M \to M, \qquad T(x) := \begin{cases} 2x - 3, & x \in [3,4] \\ 13 - 2x, & x \in [4,5] \end{cases}$$

(this is a variant of the so-called **tent map**). Since T is continuous, it is measurable. We show that $([3,5], \mathscr{B}, \mu, T)$ is a dynamical system. Since closed (or open) intervals generate the Borel σ-algebra on M, we have to show that $\mu(T^{-1}(A)) = \mu(A)$ for any closed interval $A \subset [3,5]$.

The preimage of an interval $[a,b]$ under T is

$$T^{-1}([a,b]) = \left[\frac{a+3}{2}, \frac{b+3}{2}\right] \cup \left[\frac{13-b}{2}, \frac{13-a}{2}\right].$$

This union is (up to possibly one point) disjoint. Since countable sets are null sets, we thus have

$$\begin{aligned} \mu(T^{-1}([a,b])) &= \mu\left(\left[\frac{a+3}{2}, \frac{b+3}{2}\right]\right) + \mu\left(\left[\frac{13-b}{2}, \frac{13-a}{2}\right]\right) + \mu(4) \\ &= 7 \cdot \frac{b-a}{2} + 7 \cdot \frac{b-a}{2} + 0 = 7(b-a) = \mu([a,b]), \end{aligned}$$

hence $([3,5], \mathscr{B}, \mu, T)$ is a dynamical system (not over a probability space).

This example also shows that in general $\mu(T(A)) = \mu(A)$ does not need to hold for a measure preserving function T. In our above example, we have

$$\mu(T([3,4])) = \mu([3,5]) = 14 \neq 7 = \mu([3,4]).$$

Recall that we can associate an integral to any measure space $(M, \mathscr{A}, \mu)$. We will not write down the technical definition here (see [Kle08] for details). We set

$$\mathcal{L}(M, \mathscr{A}, \mu) := \left\{ f : M \to \mathbb{R} \cup \{\pm\infty\} : f \text{ is measurable and } \int_M |f| \, \mathrm{d}\mu < \infty \right\}.$$

0.6.2 Newton's method

Newton's method is a way for approximating solutions of equations of the form $f(x) = 0$ for a differentiable function $f : \mathbb{R} \to \mathbb{R}$. In fact the method can also be adapted to systems of equations and functions $f : \mathbb{R}^n \to \mathbb{R}^n$, but we do not need

this here. Starting from an initial guess x_0 satisfying $f'(x_0) \neq 0$, the sequence (x_n) is defined iteratively by

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}. \tag{0.6.1}$$

If x_0 is not too far from a zero $\overline{x}$ with $f'(\overline{x}) \neq 0$, this is well-defined and we get (see [QSS01]):

Theorem 0.6.4 *Let $f : \mathbb{R} \to \mathbb{R}$ be continuously differentiable and $\overline{x} \in \mathbb{R}$ such that $f(\overline{x}) = 0$ and $f'(\overline{x}) \neq 0$. Then there is an $\varepsilon > 0$ such that the sequence (x_n) defined by the Newton method* (0.6.1) *converges to $\overline{x}$ for all $x_0 \in (\overline{x} - \varepsilon, \overline{x} + \varepsilon)$.*

0.6.3 Formal Languages

Given a set of symbols, we consider subsets of the free monoid generated by these symbols by means of computer science. We will do this with the notion of formal languages and finite automata, cf. [HMU02].

An **alphabet** is a finite set of symbols, usually denoted by Σ. Let Σ^* denote the set of all finite strings with symbols from Σ.

Definition 0.6.5 A **language** over some alphabet Σ is a subset of Σ^*.

We can represent languages with automata.

Definition 0.6.6 A **deterministic finite automaton** (abbreviated DFA) is a quintuple $M = (Q, \Sigma, \delta, q_0, F)$ where

- Σ is an alphabet,
- Q is a finite set, the so called **states**,
- $q_0 \in Q$ is the **start state**,
- $F \subset Q$ is a set of **final states**,
- δ is a function $\delta : Q \times \Sigma \to Q$, the **transition function**.

We can extend δ canonically to $Q \times \Sigma^*$: Let $w = a_1 \dots a_k \in \Sigma^*$ with $a_i \in \Sigma$. Then

$$\delta(q, w) := \delta(\delta(\cdots \delta(\delta(q, a_1), a_2) \cdots, a_{k-1}), a_k).$$

Mostly, automata are depicted as diagrams, cf. Figure 0.6.1. The states are shown as circles, all final states are depicted as two concentric circles. The transition function is illustrated via arrows between states q_1 and q_2, where the symbols on the arrow show the elements $s \in \Sigma$ such that $\delta(q_1, s) = q_2$. There is an additional arrow without labelling incoming in the start state.

If M is a DFA, the **language accepted by M** is defined as

$$L(M) := \{w \in \Sigma^* : \delta(q_0, w) \in F\}.$$

In some cases, we need automata with output. A **deterministic finite automaton with output** (abbreviated DFAO) is a sextuple $M = (Q, \Sigma, \delta, q_0, \Delta, \eta)$ where Q, Σ, δ, q_0 are defined as in a DFA and we further have a finite set of symbols Δ, the **output alphabet**, and an **output mapping** $\eta : Q \to \Delta$. In an illustration, we show the function η as dashed arrows and the respective values in boxes.

Note that we can turn every DFAO into a DFA by defining the set of final states F as $\eta^{-1}(D)$ for some $D \subset \Delta$.

Example 0.6.7 Let $\Sigma = \{0, 1, 2, 3, 4\}$. Then Σ^* can be viewed as the natural numbers in base 5 representation. Thus, a language over Σ is a set of natural numbers written in base 5. Consider the DFAO M given in Figure 0.6.1.

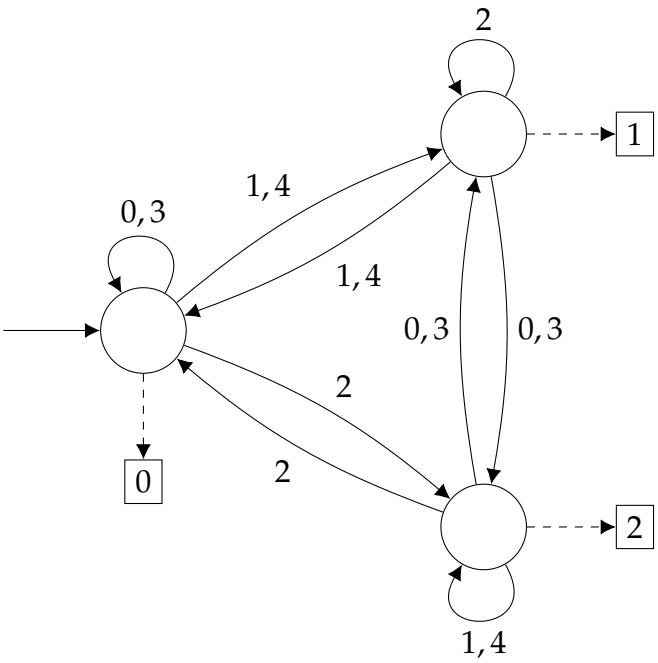

Figure 0.6.1: A deterministic finite automaton with output.

This DFAO "separates" the natural numbers in congruence classes modulo 3, i.e., we have $\eta(\delta(q_0, \langle n \rangle_5)) \equiv n \bmod 3$.

For example, $\eta(\delta(q_0, 123324130)) = 1$. Indeed, 123324130 represents the natural number 604915 and $604915 \equiv 1 \bmod 3$. If we turn this DFAO into a DFA by defining the final state q_F as the state with $\eta(q_F) = 1$, then $L(M)$ is the set of all natural numbers n written in base 5 such that $n \equiv 1 \bmod 3$.

Part I - Connections

I.1

Analysis

There are of course a lot of connections between analysis and number theory, in particular the use of complex analysis in analytic number theory. We will consider this later in Chapter I.10, here we take into account the classical part of analysis involving sequences, convergence, continuity and differentiability.

Some of the concepts presented here are in fact special cases of concepts mentioned in Paragraph 0.5.8. Since these concepts (as well as further applications thereof) belong to an important connection between number theory and analysis, and since we did not mention all details in Paragraph 0.5.8, we will develop the necessary concepts in brief for this special case instead of making use of the basics (compare the comments in Paragraph 0.5.8).

First recall one of the equivalent definitions of the real numbers $\mathbb{R}$: Let $\mathbf{c}_\infty(\mathbb{Q})$ be the ring of convergent rational sequences and $\mathbf{0}_\infty(\mathbb{Q})$ the maximal ideal (in $\mathbf{c}_\infty(\mathbb{Q})$) of sequences that converge to 0. Then one can define $\mathbb{R}$ as the quotient $\mathbf{c}_\infty(\mathbb{Q})/\mathbf{0}_\infty(\mathbb{Q})$. Here we take as notion of convergence the topology induced by the usual absolute value, which we will denote by $|\cdot|_\infty$.

Moreover, for every prime p, there is another absolute value on $\mathbb{Q}$, the p-adic absolute value $|\cdot|_p$. Write $x \in \mathbb{Q}\backslash\{0\}$ as $x = \frac{a}{b}p^r$ with $a, b, r \in \mathbb{Z}$ such that a, b and p are pairwise coprime. Then the ***p*-adic valuation** of x is defined by $v_p(x) := r$ and the ***p*-adic absolute** value is $|x|_p := p^{-r}$, $|0|_p := 0$.

As above, we can now consider convergent rational sequences with respect to the absolute value $|\cdot|_p$. We will denote this ring by $\mathbf{c}_p(\mathbb{Q})$ and its maximal ideal

of sequences converging to 0 by $\mathbf{0}_p(\mathbb{Q})$. Then the quotient $\mathbf{c}_p(\mathbb{Q})/\mathbf{0}_p(\mathbb{Q})$ is a field, the field of ***p*-adic numbers**, denoted by $\mathbb{Q}_p$.

With the p-adic numbers we can highlight some connections between number theory and analysis. First, the construction of $\mathbb{Q}_p$ is done in an analytic way (by considering convergent sequences). On the other hand, since $\mathbb{Q}_p$ is a field equipped with a norm (the unique extension of $|\cdot|_p$), we can do analysis in $\mathbb{Q}_p$. Before doing so, we mention some basic facts about the p-adic numbers. For more properties see [Neu92].

Every p-adic number x can be written uniquely in the form

$$x = \sum_{k=n}^{\infty} a_k p^k$$

with $a_k \in \{0, \ldots, p-1\}$ and $n \in \mathbb{Z}$. This is equivalent to the decimal expansion of the real numbers and is called ***p*-adic expansion**. The analogue of the ring of integers $\mathbb{Z}$ in $\mathbb{Q}$ are the ***p*-adic integers** $\mathbb{Z}_p$. These are the elements $x \in \mathbb{Q}_p$ whose p-adic expansion is of the form $\sum_{k=n}^{\infty} a_k p^k$ with $n \in \mathbb{N}_0$.

There is one important property of the p-adic absolute value $|\cdot|_p$ that makes analysis in $\mathbb{Q}_p$ somewhat different than analysis in $\mathbb{R}$ (compare Paragraph 0.5.8):

Lemma I.1.1 *Let $x, y \in \mathbb{Q}_p$. Then $|x+y|_p \leq \max\{|x|_p, |y|_p\}$.*

This is the strong triangle inequality, i.e., the p-adic absolute value is non-Archimedean. With the p-adic absolute value we can formulate the following variant of the Chinese remainder theorem, the so-called **weak approximation theorem**:

Theorem I.1.2 (weak approximation theorem) *For $n \in \mathbb{N}$ and $1 \leq i \leq n$, let $p_i \in \mathbb{P}$ be distinct primes and $a_i \in \mathbb{Q}$. Then for any $\varepsilon > 0$ there is an $x \in \mathbb{Q}$ with $|x - a_i|_{p_i} < \varepsilon$.*

In the p-adic absolute value, small distances correspond to high divisibility by p, i.e., the weak approximation theorem states that for any $k \in \mathbb{N}$ there is an $x \in \mathbb{Q}$ with $x \equiv a_i \bmod p_i^k$.

The approximation theorem holds true in a slightly more general version, see [Neu92]. In Chapter I.4 we will even see a generalization thereof, the strong approximation theorem.

After reviewing these basic facts we wish to solve polynomial equations in $\mathbb{Z}_p$. This can be done with **Hensel's lemma**. There are many versions of Hensel's lemma (see [Neu92, Conc]), its basic form is the following theorem.

Theorem I.1.3 (Hensel's lemma) *Let $f \in \mathbb{Z}[x]$ and $a \in \mathbb{Z}$ with*

$$f(a) \equiv 0 \bmod p \qquad \text{and} \qquad f'(a) \not\equiv 0 \bmod p.$$

Then there is a unique $\alpha \in \mathbb{Z}_p$ with $f(\alpha) = 0$ and $\alpha \equiv a \bmod p$.

Hensel's lemma itself is another use of analysis in number theory, since (one of) its proof is essentially Newton's method, cf. Chapter I.3.

Hensel's lemma has a lot of applications. With its help, one can find squares in $\mathbb{Z}_p$ (see [Conc]), square roots in $\mathbb{Q}_p$ (see [Rob99]) and roots of unity in $\mathbb{Q}_p$ (see [Conc, Rob99]). A stronger version of Hensel's lemma can also be used to show that, in contrast to the field $\mathbb{R}$, the p-adic numbers $\mathbb{Q}_p$ cannot be ordered. This can be shown by representing 0 as a sum of nonzero squares using the four-squares theorem (Theorem 0.5.12).

Breaking Hensel's lemma down to its main statement we can (under certain conditions on the polynomial f) lift solutions of $f(x) \equiv 0 \bmod p$ to roots of f in $\mathbb{Z}_p$, i.e., $f(x) = 0$ is solvable in $\mathbb{Z}_p$ if and only if $f(x) \equiv 0 \bmod p$ is solvable.

Going one step further, we take a look at the **Hasse principle**, an important **local-global principle** in number theory. The idea of the Hasse principle is the following:

Given a polynomial $f \in \mathbb{Z}[x_1, \ldots, x_n]$, we want $f(\mathbf{x}) = 0$ to be solvable in $\mathbb{Q}$ if and only if $f(\mathbf{x}) = 0$ is solvable in $\mathbb{R}$ and in $\mathbb{Q}_p$ for all primes p. We say that the Hasse principle is true if this is possible.

The classical case where the Hasse principle is true is the **Hasse-Minkowski theorem** about quadratic forms, see [Sch07a, Ser73] for a direct approach and [O'M73] for a more sophisticated version (with the use of ideles, a tool that we will introduce in Chapter I.4).

Theorem I.1.4 (Hasse-Minkowski theorem) *Let f be a rational quadratic form. Then $f(\boldsymbol{x}) = 0$ is solvable in $\mathbb{Q}$ if and only if $f(\boldsymbol{x})$ is solvable in $\mathbb{R}$ and in $\mathbb{Q}_p$ for all p.*

Unfortunately, the Hasse principle is not true for arbitrary polynomials f (not even if f is homogeneous). The two standard counterexamples are **Selmer's cubic** $f(x, y, z) = 3x^3 + 4y^3 + 5z^3$ (see [Sel51, Cong]) and the equation of **Lind and Reichardt** $f(x, y) = x^4 - 2y^2 - 17$ (see [Sch07a]). These polynomials have roots in $\mathbb{R}$ and in $\mathbb{Q}_p$ for all p, but no rational roots. More counterexamples can be found in [AL11].

There are still many open questions related to the Hasse principle and the cases in which it is true. It is also possible to attack the Hasse principle with tools from analytic number theory, cf. Chapter I.10. For a list of other local-global principles in mathematics see [qua].

Let's turn our attention to analysis in the p-adic numbers. Since $\mathbb{Q}_p$ is a field equipped with a norm, we can define the notion of continuous and differentiable functions in the usual way. Thus a function f (defined on a suitable subset U of $\mathbb{Q}_p$) is differentiable at $a \in U$ if the limit $\lim_{x\to a} \left|\frac{f(x)-f(a)}{x-a}\right|_p$ exists. As usual we will denote the derivative of f at a point a by $f'(a)$.

Some of the properties of differentiable functions are rather different to those of real differentiable functions. For example, consider the function $f : \mathbb{Z}_p \to \mathbb{Z}_p$ defined by

$$f(x) := \begin{cases} p^{2n}, & \text{if } x \equiv p^n \bmod p^{2n+1} \\ 0, & \text{otherwise} \end{cases}.$$

Then f is not locally constant at $x = 0$ (since $f(p^n) = p^{2n}$ and $p^n \to 0$), but f is differentiable with $f'(a) = 0$ for all $a \in \mathbb{Z}_p$. Hence p-adic functions with derivative 0 need not be constant. This means that in p-adic analysis the usual notion of differentiability is not useful. Instead, one requires **strictly differentiable** functions, cf. [Rob99]. More examples of functions with counterintuitive behaviour can be found in [Kat07].

We examine two special functions. As in the real numbers we can define the exponential function $\exp(x)$ and the logarithm $\log(x)$ via its series

$$\exp(x) := \sum_{k=1}^{\infty} \frac{x^k}{k!} \qquad \text{and} \qquad \log(x+1) := \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^k}{k}.$$

While the exponential series is everywhere convergent in the real numbers, this is not true if we view it as a series in $\mathbb{Q}_p$. In fact we have (see [Rob99]):

Theorem I.1.5 *Let* $\exp(x)$ *and* $\log(x+1)$ *be defined by the series representations above.*

- *The series representation for* $\exp(x)$ *converges if and only if* $|x|_p < p^{-\frac{1}{p-1}}$.
- *The series representation for* $\log(x+1)$ *converges if and only if* $|x|_p < 1$.

Hence, there are $x \in \mathbb{Q}_p$ such that $\exp(x)$ is not defined. Other results about the functions exp and log (partially similar to the real case) can be found in [Rob99]. In particular, the same functional equations hold, as long as all series involved converge.

There is one thing we want to draw attention to: If the exponential function is defined over the real numbers and extended to the complex numbers, there is a nonzero element $z \in \mathbb{C}$ (namely $2\pi i$) such that $\exp(z) = 1$, which means, together with the functional equation for the exponential function, that this function has a period. We examine this for the p-adic exponential function. We consider exp as a function on $\mathbb{C}_p$, where $\mathbb{C}_p$ is the completion of the algebraic closure of $\mathbb{Q}_p$ (compare [Rob99]). Clearly the number π makes no sense in $\mathbb{C}_p$, so we have to search for another element z. But in fact there is no nonzero $z \in \mathbb{C}_p$ such that $\exp(z) = 1$. This is a consequence of **Strassmann's theorem**, see [Rob99, Kat07]:

Theorem I.1.6 (Strassmann's theorem) *Let $A \subset \mathbb{C}_p$ and let f be a nonzero power series with coefficients $a_k \in A$. If $|a_k|_p \to 0$, $f(z)$ has only finitely many zeros in A.*

When applied to $f(z) = \exp(pz) - 1$, Strassmann's theorem shows that there are only finitely many zeros of $f(z)$ in $\mathbb{C}_p$, thus there is no nonzero $z \in \mathbb{C}_p$ with $\exp(z) = 1$ (otherwise the functional equation for $\exp(z)$ would yield infinitely many zeros of $f(z)$). Hence, the p-adic exponential and logarithm do not share all the nice properties of the respective complex functions.

But what if we want a formula of some kind like $\exp(2\pi i) = 1$ to be true in a p-adic sense? We have seen that this cannot hold in $\mathbb{Q}_p$, so we briefly consider another setting. For a more detailed exposition see [BP15] or [Gos96].

Let q be a prime power and $k = \mathbb{F}_q(T)$ be the field of rational functions over $\mathbb{F}_q$ in one variable. We consider the following absolute value $|\cdot|_T$ on k (this is completely analogous to the p-adic absolute value): Write $f \in k$ in the form $T^e \frac{g}{h}$ such that g, h and T are pairwise coprime. Then $|f|_T := q^{-e}$. (In the literature, this absolute value is usually denoted by $|\cdot|_\infty$. Since we used this notation already for the usual absolute value on $\mathbb{Q}$, we will write $|\cdot|_T$.) Let $\mathbb{C}_\infty$ be the completion of the algebraic closure of the completion of k with respect to $|\cdot|_T$. Although this construction seems a bit confusing, this is in fact completely analogous to the construction of $\mathbb{C}$ in the real case and $\mathbb{C}_p$ in the p-adic case, cf. Table I.1.1.

base field	absolute value	completion	algebraic closure of the completion	completion of the algebraic closure of the completion
$\mathbb{Q}$	$\lvert\cdot\rvert_\infty$	$\mathbb{R}$	$\mathbb{C}$	$\mathbb{C}$
$\mathbb{Q}$	$\lvert\cdot\rvert_p$	$\mathbb{Q}_p$	$\overline{\mathbb{Q}_p}$	$\mathbb{C}_p$
$\mathbb{F}_q(T)$	$\lvert\cdot\rvert_T$	k_∞	$\overline{k_\infty}$	$\mathbb{C}_\infty$

Table I.1.1: Comparison of $\mathbb{C}$, $\mathbb{C}_p$, and $\mathbb{C}_\infty$.

We can construct an exponential function $e_C : \mathbb{C}_\infty \to \mathbb{C}_\infty$, the **Carlitz exponential** (compare [Gos96]), such that the smallest period of e_C is $\widetilde{\pi}_q = (T - T^q)^{\frac{1}{q-1}} \pi_q$ with

$$\pi_q := \prod_{k=1}^{\infty} \left(1 - \frac{T^{q^k} - T}{T^{q^{k+1}} - T} \right).$$

We call π_q **Carlitz π**. Thus, in $\mathbb{C}_\infty$ we have an anlogue to π. The number π_q shares even more properties with the real number π, for example it occurs at special values of the **Carlitz zeta function**, cf. [Car35]. One more similarity is that π_q is transcendental over $\mathbb{F}_q(T)$. This has first been shown in [Wad41]. We will see an interesting way to prove this in Chapter I.11. It has also been shown that other special values of the Carlitz zeta function are transcendental, see [DH91].

Another connection between analysis and number theory will be shown in Chapter II.4. We will consider the solvability of a puzzle that can be modeled with linear algebra. To examine the solvability one can use methods from (algebraic) number theory as well as methods from analysis, namely improper Riemann integrals.

I.2

Sequences

There are a lot of interesting sequences arising from number theoretic questions (for example Beatty sequences) as well as many number theoretic questions concerning sequences, in particular if the sequences are integer valued (for example squares in the Fibonacci sequence, cf. [Coh64]). Here we take a look at the **Farey sequence**.

Definition I.2.1 Let $n \in \mathbb{N}$. The Farey sequence of order n ist the finite sequence $\mathscr{F}_n$ consisting of all fractions $\frac{a}{b}$ in increasing order with $a \in \mathbb{N}_0, b \in \mathbb{N}, a \leq b \leq n$ and $\gcd(a, b) = 1$.

In other words, the Farey sequence $\mathscr{F}_n$ consists of all reduced fractions α in the intervall $[0, 1]$ whose denominator is less or equal to n. An immediate consequence is that the number of elements in $\mathscr{F}_n$ is $1 + \sum_{k=1}^{n} \varphi(k)$. For example, we have

$$\mathscr{F}_6 = \left(\frac{0}{1}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{1}{1}\right).$$

The Farey sequence has a lot of interesting properties, most of which concerning neighbour fractions, i.e., two fractions $\alpha, \beta \in \mathscr{F}_n$ that are direct successors. Let $\alpha = \frac{a_1}{b_1}, \beta = \frac{a_2}{b_2}, \gamma = \frac{a_3}{b_3}$ be three succeeding fractions in $\mathscr{F}_n$. Then we have the following properties (compare [SF07, HW08]).

- β is the mediant of α and γ, i.e., $\beta = \frac{a_1+a_3}{b_1+b_3}$ (this representation need not be reduced).
- We have $b_1a_2 - a_1b_2 = 1$. The opposite of this statement is also true: If $\frac{a_1}{b_1}$ and $\frac{a_2}{b_2}$ are two reduced fractions, both less than 1, with $b_1a_2 - a_1b_2 = 1$, then the two fractions are neighbours in $\mathscr{F}_n$ for all n with $\max(b_1, b_2) \leq n < b_1 + b_2$.
- The inequality $b_1 + b_2 > n$ holds.

Some neighbour fractions also have related continued fraction expansions. Let $\langle 0; a_1, \ldots, a_{m-1}, 1\rangle$ be the (not normalized) continued fraction expansion of $\beta = \frac{a}{n}$. To avoid special cases, we suppose $n \geq 3$. Let α, γ be the neighbours of β in $\mathscr{F}_n$. Since fractions with the same denominator $k \geq 2$ cannot be neighbours in any Farey sequence, α and γ have denominator smaller than n. Further, they do not have the same denominator, since α and γ are neighbours in $\mathscr{F}_{n-1}$. Let α have larger denominator than γ. Then the continued fraction expansion of α is $\langle 0; a_1, \ldots, a_{m-1}\rangle$ and that of γ is $\langle 0; a_1, \ldots, a_{m-2}\rangle$. Using the second point from the above list, this follows from the recursion formula for the continued fraction expansion. Note that these expansions do not hold if the initial fraction β has denominator strictly less than n or if we consider neighbours in the Farey sequence $\mathscr{F}_k$ with $k \neq n$.

Apart from having interesting properties on its own, Farey sequences serve as a useful tool for number theoretic questions. One of the most prominent occurrences of Farey sequences is in the proof of **Dirichlet's approximation theorem** in its weak form:

Theorem I.2.2 (Dirichlet's approximation theorem) *Let $\alpha \in \mathbb{R}, n \in \mathbb{N}$. Then there is a fraction $\frac{a}{b}$ with $0 < b \leq n$ such that*

$$\left|\alpha - \frac{a}{b}\right| \leq \frac{1}{b(n+1)}.$$

For its proof (see [SF07]) one essentially uses the three facts from above.

Besides this (somehow direct) use of Farey fractions, the Farey sequence gives an alternative formulation for the Riemann hypothesis. To get this, let $\widetilde{\mathscr{F}_n}$ be the Farey sequence without the fraction $\frac{0}{1}$. Let further A be the number of positive reduced fractions with denominator less or equal to n (i.e., the number of elements in $\widetilde{\mathscr{F}_n}$) and let ρ_ν ($\nu = 1, \ldots, A$) run through all values of $\widetilde{\mathscr{F}_n}$ in incresaing order. Set $\eta_\nu = \rho_\nu - \frac{\nu}{A}$. Then we have the following theorem (see [Fra24, Lan24]).

Theorem I.2.3 *The Riemann hypothesis is equivalent to each of the following statements:*

1. *For all $\varepsilon > 0$, we have*

$$\sum_{\nu=1}^{A} \eta_\nu^2 = \mathcal{O}(n^{-1+\varepsilon}).$$

2. *For all $\varepsilon > 0$, we have*

$$\sum_{\nu=1}^{A} |\eta_\nu| = \mathcal{O}(n^{\frac{1}{2}+\varepsilon}).$$

The reason for this lies in the formula

$$M(n) = \sum_{r=1}^{n} \mu(r) = \sum_{\nu=1}^{A} e^{2\pi i \rho_\nu}$$

and in the formulation of the Riemann hypothesis involving the **Mertens function** $M(n)$ (here $\mu(n)$ denotes the **Möbius function**). For a discussion and more equivalent formulations of the Riemann hypothesis involving Farey sequences see [KY96]. We will see one more number theoretic application of Farey sequences in Chapter I.10.

Farey sequences also have applications outside mathematics, in particular in physics. Since Chapter I.11 about physics and computer science is devoted to a completely different connection between physics and number theory, we will briefly mention some application of Farey sequences in physics here.

In [Tom14], the author points out connections between Farey sequences and resonance diagramms. A relationship between the Farey sequence and circuits can be found in [Mar12], and in [DHPB15] the authors use Farey sequences for digital image processing.

At the end of this chapter we take a quick look at geometric objects closely related to the Farey sequence, the so called **Ford circles**:

Definition I.2.4 Let $a, b \in \mathbb{N}$ be coprime with $a < b$. The Ford circle $C[a/b]$ is the circle with center at $\left(\frac{a}{b}, \frac{1}{2b^2}\right)$ and radius $\frac{1}{2b^2}$.

One usually also denotes half of the circles $C[0/1]$ and $C[1/1]$ as Ford circles. The Ford circles $C[a/b]$ for $b \leq 6$ can be found in Figure I.2.1.

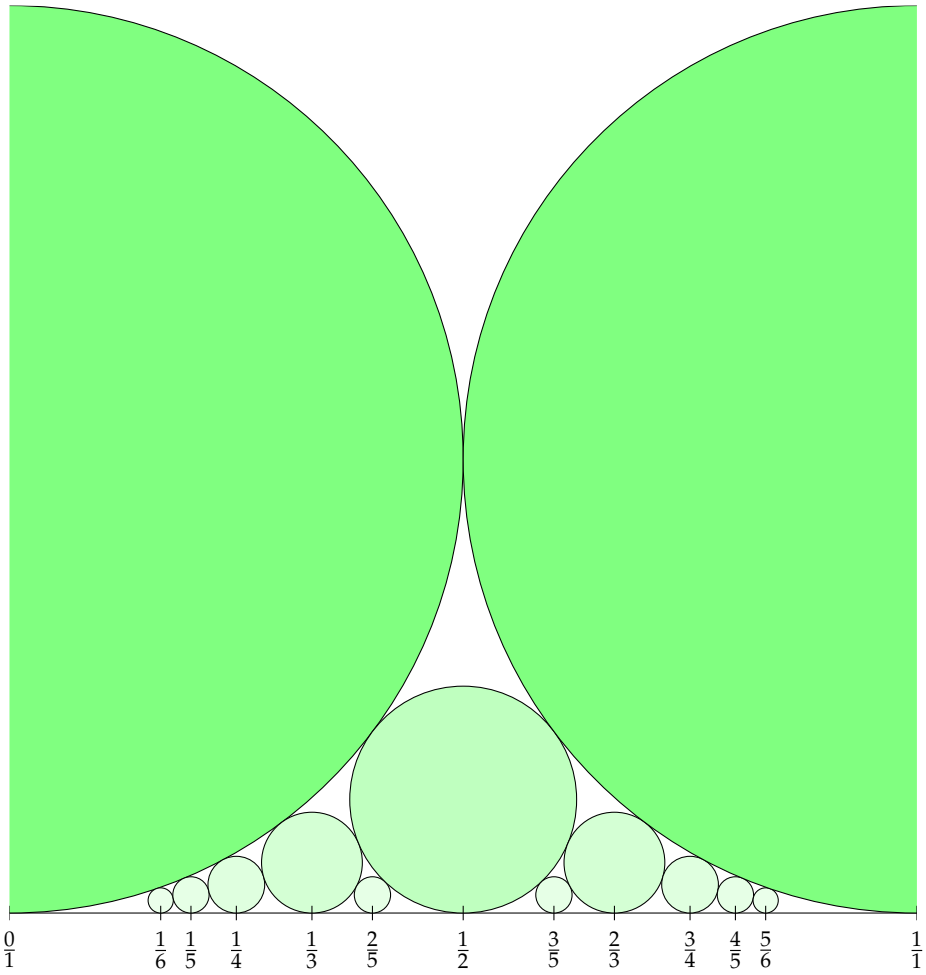

Figure I.2.1: The Ford circles $C[a/b]$ for $b \leq 6$.

Two Ford circles $C[a_1/b_1]$ and $C[a_2/b_2]$ are either disjoint or tangent, where the second case occurs if and only if $\frac{a_1}{b_1}$ and $\frac{a_2}{b_2}$ are neighbours in a Farey sequence. We will see a nice application of Ford circles in Chapter I.9.

I.3

Applied Mathematics

In this chapter we discuss connections between number theory and probability theory, respectively numerical analysis.

Concerning probability theory there are two main aspects: probabilistic number theory and the use of ergodic theory in number theory. In probabilistic number theory, we have, among others, the classical results of Erdős-Wintner, cf. [EW39], and Erdős-Kac, cf. [EK40]. The latter is a consequence of the fact that for a given number n and two primes p_1, p_2 the events "n is divisible by p_i" are independent, i.e., in some sense primes behave like independent random variables. For more about probabilistic number theory we refer to [Ten07].

Here we focus on the use of ergodic transformations in number theory.

Definition I.3.1 Let $(M, \mathscr{A}, \mu, T)$ be a dynamical system over a probability space. We call T **ergodic** (with respect to μ) if for all $A \in \mathscr{A}$ the equality $T^{-1}(A) = A$ implies either $\mu(A) = 0$ or $\mu(A) = 1$.

In other words, T is called ergodic if all measurable sets A that are invariant under T^{-1} have measure 0 or 1. There are more equivalent formulations of ergodic transformations, see [EW11, Möl, Ste]. Note that T need not be ergodic if $\mu(A) = 0$ or $\mu(A) = 1$ holds for every A with $T(A) = A$, cf. Example 0.6.3.

One of the most important theorems in ergodic theory is Birkhoff's pointwise ergodic theorem, see [Möl, Ste].

Theorem I.3.2 (Birkhoff's pointwise ergodic theorem) *Let $(M, \mathscr{A}, \mu, T)$ be a dynamical system over a probability space and $f \in \mathcal{L}(M, \mathscr{A}, \mu)$. Then for almost all $y \in M$ the limit $f^*(y) := \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} f(T^k y)$ exists and we have*

- $f^* \in \mathcal{L}(M, \mathscr{A}, \mu)$,
- $f^*(Ty) = f^*(y)$,
- $\int_M f^* \, \mathrm{d}\mu = \int_M f \, \mathrm{d}\mu$.

If further T is ergodic, f^* is constant almost everywhere and

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} f(T^k y) = f^*(y) = \int_M f \, \mathrm{d}\mu.$$

We can interpret $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} f(T^k y)$ as the time mean of f and $\int_M f \, \mathrm{d}\mu$ as the space mean of f, varying over all of M. Birkhoff's theorem tells us that these two means coincide if T is ergodic. This fact is the one which is most important for our purpose here.

Now we want to see some applications of Birkhoff's theorem to ergodic transformations that occur in number theory. To use this theorem appropriately, we would have to show that the transformations we are considering are ergodic. This is in most cases too technical, thus we will not do this here. For this proofs we refer to [Möl, Ste].

Our main focus is on infinite continued fractions. Of course it suffices to study continued fractions of real numbers x with $0 \leq x < 1$, since other numbers would only change the first partial quotient.

Let T be the transformation

$$T : [0,1) \to [0,1), \quad T(x) := \begin{cases} \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor, & x \neq 0 \\ 0, & x = 0 \end{cases}.$$

Directly from the construction of continued fractions we get

$$T\left(\langle 0; a_1, a_2, \ldots \rangle\right) = \langle a_1; a_2, a_3, \ldots \rangle - \left\lfloor \langle a_1; a_2, a_3, \ldots \rangle \right\rfloor = \langle 0; a_2, a_3, \ldots \rangle.$$

To use Birkhoff's theorem, we need a probability space $([0,1), \mathscr{A}, \mu)$ such that μ is T-invariant. In fact $([0,1), \mathscr{B}, \mu, T)$ is a dynamical system over a probability space, where

$$\mu(A) := \frac{1}{\log 2} \int_A \frac{1}{1+x} \, \mathrm{d}x$$

and moreover, T is ergodic. These statements even remain true when replacing $\mathscr{B}$ by the bigger σ-algebra of Lebesgue measurable sets, but we do not need this here. Now one can prove the theorems of **Khintchine** and **Lévy**. Proofs of both theorems can be found in [EW11, Möl, Ste].

Theorem I.3.3 (Khintchine's theorem) *For almost all $x = \langle 0; a_1, a_2, \ldots \rangle \in [0,1)$ we have:*

1. *The number $k \in \mathbb{N}$ appears in the continued fraction expansion with asymptotic density*
$$\lim_{n\to\infty} \frac{1}{n} \left| \{1 \le j \le n : a_j = k\} \right| = \frac{1}{\log 2} \log\left(1 + \frac{1}{k(k+2)}\right).$$

2. *The arithmetic mean of the partial quotients is*
$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} a_k = \infty.$$

3. *The geometric mean of the partial quotients is*
$$\lim_{n\to\infty} \left(\prod_{k=1}^{n} a_k \right)^{\frac{1}{n}} = \prod_{k=1}^{\infty} \left(1 + \frac{1}{k(k+2)}\right)^{\frac{\log k}{\log 2}} \approx 2.6854520010.$$

The first statement can be proved with the indicator function $f = \chi_{(\frac{1}{k+1}, \frac{1}{k}]}$, since $a_n = k$ if and only if $T^{n-1}(x) \in (\frac{1}{k+1}, \frac{1}{k}]$. Thus we get

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} \left| \{1 \le j \le n : a_j = k\} \right| &= \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \chi_{(\frac{1}{k+1}, \frac{1}{k}]}(T^{j-1}x) \\
&= \int_0^1 \chi_{(\frac{1}{k+1}, \frac{1}{k}]} \, \mathrm{d}\mu \\
&= \frac{1}{\log 2} \int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{\mathrm{d}x}{1+x} \\
&= \frac{1}{\log 2} \left(\log\left(1 + \frac{1}{k}\right) - \log\left(1 + \frac{1}{k+1}\right) \right) \\
&= \frac{1}{\log 2} \log\left(1 + \frac{1}{k(k+2)}\right).
\end{aligned}$$

For the second statement we can proceed analogously by using the step function $f(x) = k$ if $x \in (\frac{1}{k+1}, \frac{1}{k}]$ and for the third statement we consider the step function $f(x) = \log k$ if $x \in (\frac{1}{k+1}, \frac{1}{k}]$.

Theorem I.3.4 (Lévy's theorem) *For $x \in [0,1)$ let $\frac{p_n}{q_n}$ be the n-th convergent of x. Then for almost all $x = \langle 0; a_1, a_2, \ldots \rangle \in [0,1)$ we have:*

1. $\lim_{n\to\infty} \frac{1}{n} \log(q_n) = \frac{\pi^2}{12 \log 2}$.
2. $\lim_{n\to\infty} \frac{1}{n} \log \left| x - \frac{p_n}{q_n} \right| = -\frac{\pi^2}{6 \log 2}$.

It is clear that these theorems cannot hold for all $x \in [0,1)$. In particular, all five statements are wrong for any rational x, since the statements can only hold for infinite continued fractions. If x is of the form $a + b\sqrt{d}$ with $a, b \in \mathbb{Q}$ and $d \in \mathbb{Z}\backslash\{1\}$ squarefree, then we know from Theorem 0.5.8 that the partial quotients are bounded. This immediately yields that the first two statements in Khintchine's theorem cannot be true. Further, the geometric mean of such a continued fraction equals the geometric mean of the period of this fraction. Thus, the third statement can only be true if the **Khintchine constant**

$$K := \prod_{k=1}^{\infty} \left(1 + \frac{1}{k(k+2)}\right)^{\frac{\log k}{\log 2}} \approx 2.6854520010$$

is algebraic. This is not known yet. In fact it is unkown if K is irrational.

If Φ denotes the **golden ratio**, i.e., $\Phi = \frac{1+\sqrt{5}}{2}$, then p_n and q_n are consecutive Fibonacci numbers, and using the explicit form of Fibonacci numbers it is easy to show that $\lim_{n\to\infty} \frac{1}{n} \log(q_n) = \log(\Phi)$, thus the first statement in Lévy's theorem does not hold for Φ. For numbers x with bounded partial quotients we have

$$\lim_{n\to\infty} \frac{1}{n} \log(q_n) < \lim_{n\to\infty} \frac{1}{n} \log((B+1)^n) = B + 1 < \infty$$

if B is the biggest partial quotient of x, such that this part is not too far away from the truth for suitable B (and can possibly be true for some x of this form). On the other hand, if x is algebraic of degree 2, i.e., $x = a + b\sqrt{d}$ with $a, b \in \mathbb{Q}, d \in \mathbb{Z}\backslash\{1\}$ squarefree, we get (analogously to the case $x = \Phi$) $\lim_{n\to\infty} \frac{1}{n} \log(q_n) = \log(\alpha)$ for some algebraic α. Thus the first statement in Lévy's theorem can only hold for (some) algebraic x of degree 2 if $e^{\frac{\pi^2}{12\log(2)}}$ is algebraic. This is (as far as the author knows) an open question, but there is no reason to believe that $e^{\frac{\pi^2}{12\log(2)}}$ is algebraic. Hence the first statement in Lévy's theorem is probably false for all algebraic x of degree 2.

Concerning the second statement in Lévy's theorem, Liouville's theorem (Theorem 0.5.20) immediately yields

$$\lim_{n\to\infty}\frac{1}{n}\log\left|x-\frac{p_n}{q_n}\right|=0$$

for algebraic x (of arbitrary degree). Hence the second statement in Lévy's theorem cannot be true for algebraic numbers. We will see another counterexample to the second statement in Lévy's theorem in Chapter II.4.

Finally we consider a transcendental number whose continued fraction expansion is well known, namely $e=\langle 2;1,2,1,1,4,1,1,6,1,1,\ldots\rangle$. We see that 1 occurs with asymptotic density $\frac{2}{3}$ instead of $\frac{\log(4/3)}{\log(2)}\approx 0.415$. This means that the first statement in Khintchine's theorem does not hold. The second statement does hold, since the arithmetic mean is asymptotically $\frac{1}{9}n$. The third statement again does not hold, since the geometric mean is asymptotically $\left(\frac{2n}{3e}\right)^{\frac{1}{3}}$.

These examples are somewhat typical: There are only very few numbers, for which we know that (parts of) the statements in the theorems of Khintchine and Lévy are true, despite the fact, that they are true for almost every number. For numbers like π and the Khintchine constant K, it is believed that the statements hold, but this is still unproven.

Theorem I.3.3 shows how ergodic theory can be used to get results about certain distributions, in our case about digits in continued fractions. We can also use ergodic theory to get results about uniform distribution. Recall that a sequence $(x_n)\subset[0,1)$ is called **uniformly distributed** if

$$\lim_{n\to\infty}\frac{1}{n}\left|\{k:1\le k\le n, x_k\in[a,b]\}\right|=b-a$$

for all $0\le a<b<1$.

To see how this relates to ergodic theory, and in particular Birkhoff's ergodic theorem, we mention a result of Weyl (see [Ste]):

Theorem I.3.5 *A sequence $(x_n)\subset[0,1)$ is uniformly distributed if and only if for any Riemann-integrable function $f:[0,1]\to\mathbb{C}$ we have*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}f(x_n)=\int_0^1 f(x)\,\mathrm{d}x.$$

Here we have again some kind of comparison between space means and time means.

There is another criterion, also due to Weyl, for uniform distribution (see [Ste]):

Theorem I.3.6 *A sequence $(x_n) \subset [0,1)$ is uniformly distributed if and only if*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} e(mx_n) = 0$$

for any integer $m \neq 0$. Here $e(x) = e^{2\pi i x}$.

This criterion can be used to show that the sequence $(\{n\xi\})_n$ (here $\{x\}$ denotes the fractional part of x) is uniformly distributed if and only if ξ is irrational. We do not need any ergodic theory for this result, but it can be generalized:

Theorem I.3.7 *Let $f = a_n x^n + \cdots + a_0 \in \mathbb{R}[x]$ a polynomial such that at least one of the coefficients a_i is irrational. Then the sequence $(\{P(n)\})_n$ of the fractional parts is uniformly distributed.*

For a proof using ergodic theory see [Möl]. For even more applications of ergodic theory to number theory (for example to the Riemann ζ-function) and other mathematical topics see [Möl, Ste, Ste12].

We conclude this chapter with some brief remarks about connections between number theory and numerical analysis.

One important topic in numerical analysis is the numerical evaluation of integrals. If $f : [0,1] \to \mathbb{C}$ is some Riemann-integrable function, we could choose points $x_k \in [0,1]$ and try to approximate the integral $\int_0^1 f(x)\,\mathrm{d}x$ with the sum $\frac{1}{n}\sum_{k=1}^{n} f(x_k)$. The error in this estimation will, among others, depend on the choice of the points x_k. In this chapter we have already seen how we should choose the points x_k. Theorem I.3.5 about uniform distribution tells us that we can numerically integrate a function f with relatively small error if the sample points are well distributed (which is not surprising). One can also use methods from elementary number theory to explicitely construct sets of points for which the approximation is good (even for higher-dimensional integrals), see [Pil].

On the other hand, we already used a method from numerical analysis in Chapter I.1: Hensel's lemma can be proved using iteration methods like Newton's method. Such iterative methods can also be used to compute inverses of an element a modulo a prime power p^s if the inverse of a modulo p is known, see [KX10].

I.4

Topology

Two of the most prominent uses of topology in number theory are **étale cohomology** and ***K*-theory**. Since these concepts require too much prior knowledge to explain and understand, we will just mention the kind of problems that one tries to solve with them and we will not explain all necessary notation. Instead we will highlight another use of topology in number theory.

The main idea behind étale cohomology is that "Counting points over a finite field is asking for points fixed by Frobenius, and Weil's dream was to derive a formula by some (at the time) unknown analogue of the Lefschetz fixed point formula in algebraic topology. In a nutshell that is what first brought cohomology into arithmetic, for a very specific purpose" (K. Conrad in [Conj]), see also [Oss]. For an introduction to étale cohomology we refer to [Mil13].

The main application of étale cohomology in number theory are the **Weil conjectures** (in fact these are proven and thus no conjectures anymore). We will roughly state them without explaining all terminology.

The Weil conjectures came up in a paper of Weil in 1949 (see [Wei49]). Let q be a prime power and X be a projective algebraic variety over $\mathbb{F}_q$. Let N_m denote the number of points of X over $\mathbb{F}_{q^m}$. Then we can attach a zeta function $\zeta(X,s)$ to X, defined by

$$\zeta(X,s) = \exp\left(\sum_{m=1}^{\infty} \frac{N_m}{m} q^{-ms}\right).$$

The Weil conjectures are four conjectures about the function $\zeta(X, s)$, motivated by the Riemann ζ-function. The basic statements of the first three conjectures are:

1. $\zeta(X, s)$ is a rational function in q^{-s}.
2. $\zeta(X, s)$ has a functional equation, i.e., we can relate $\zeta(X, s)$ with $\zeta(X, n - s)$ (here n is the dimension of X).
3. The Riemann hypothesis holds for $\zeta(X, s)$. This can be formulated via a condition on the zeros of the polynomials appearing in the representation of $\zeta(X, s)$ as a rational function.

We will not mention the fourth conjecture here, since this would need to much terminology, we refer to [Oss].

Dwork proved the first conjecture in 1960 (see [Dwo60]), the second and fourth conjecture were proved by Grothendieck in 1965 (see [Gro65]). The proof of the third conjecture was due to Deligne in 1974 (see [Del74]).

Étale cohomology can also be used to establish similarities of number theoretic objects (primes) and topological objects (knots), see [LS12].

K-theory deals with groups $K_m(C)$ associated to a certain object C. There are two main objects one considers, namely projective modules over rings (this is called **algebraic K-theory**) and vector bundles over a topological space (this is called **topological K-theory**). For the easiest cases, i.e., the group $K_0(C)$, one turns the abelian monoid of projective modules, respectively the abelian monoid of vector bundles, into a group via a certain equivalence relation on its cartesian product (similar to the construction of $\mathbb{Z}$ from $\mathbb{N}$). If the underlying topological space (M, τ) of the vector bundles is compact and Hausdorff, the corresponding group is isomorphic to the K-group associated to the ring of continuos functions on (M, τ) with complex values, see [Fri07]. For higher m, the groups $K_m(C)$ are defined via fundamental groups.

The groups $K_m(C)$ carry important structural information about the object C. If K is a number field, then $K_0(\mathcal{O}_K) = \mathbb{Z} \oplus \mathcal{C}l_K$ and $K_1(\mathcal{O}_K) = \mathcal{O}_K^*$, see [Bes14]. Even for the easiest case $C = \mathbb{Z}$, only the groups $K_m(\mathbb{Z})$ for $m \leq 4$ are known. For higher m only some structure is known, see [Gha99].

From a more conceptual viewpoint, K-theory is a functor between some categories and the category of abelian groups. The groups $K_m(C)$ can coincide with groups obtained via different constructions such as (étale) cohomology or representation theory, see [Kuk] and the sources mentioned there.

An important application of K-theory in algebraic number theory is the **Kummer-Vandiver conjecture**. This is in fact still an open problem.

Conjecture I.4.1 (Kummer-Vandiver conjecture) *Let p be a prime, ζ_p a primitive p-th root of unity and $\mathbb{Q}(\zeta_p)^{\mathbb{R}}$ be the maximal totally real subfield of $\mathbb{Q}(\zeta_p)$. Then p does not divide the class number of $\mathbb{Q}(\zeta_p)^{\mathbb{R}}$.*

For an introduction to the Kummer-Vandiver conjecture and the use of K-theory in the attempts to prove it, see [Gha99]. For this conjecture one needs the groups $K_{2n-2}(\mathbb{Z})$ (for some n). Under the assumption of some conjectures more is known about these groups and the structure will depend on values of the Riemann ζ-function, cf. [Sou08]. The Kummer-Vandiver conjecture can be used for proving Fermat's last theorem, cf. Theorem I.8.1.

Let us now consider a case where a topological definition yields a useful number theoretic object (for proofs see [Cas67]). Let K be a **global field**, i.e., a finite seperable extension of $\mathbb{Q}$ or $\mathbb{F}_q(t)$ for some prime power q. For any valuation ρ on K let K_ρ be the completion of K with respect to ρ. If ρ is non-Archimedean, let $\mathcal{O}_\rho$ be the **valuation ring** of K_ρ, i.e.,

$$\mathcal{O}_\rho = \{x \in K_\rho : \rho(x) \geq 0\}.$$

These rings are compact, see [Neu92]. The set of **finite adeles** is defined as the restricted topological product (over all non-Archimedean ρ) of the fields K_ρ with respect to the rings $\mathcal{O}_\rho$. The **adele ring** V_K of K is defined as the cartesian product of the set of finite adeles with all completions K_ρ such that ρ is Archimedean.

Elements of V_K are called **adeles**. By defining multiplication and addition on V_K componentwise this gives a ring. With the topology defined in Paragraph 0.1.1 V_K becomes a topological ring.

As already indicated, we will see a generalization of the variant of the Chinese remainder theorem we have seen in Chapter I.1. This is the **strong approximation theorem**, see [Cas67]:

Theorem I.4.2 (strong approximation theorem) *Let K be a global field and ρ_0 be any valuation of K. Let $V_K^{\rho_0}$ be the restricted topological product over all $\rho \neq \rho_0$ of the fields K_ρ with respect to the rings $\mathcal{O}_\rho$. Then the image of the canonical embedding of K is dense in $V_K^{\rho_0}$.*

This is not true if we consider the complete ring of adeles. In fact, K is discrete in V_K.

Now we take a look at invertible adeles, i.e., adeles $x \in V_K$ such that there is an adele $y \in V_K$ with $x \cdot y = (1, 1, 1, \ldots)$. Let J_K be the subset of invertible elements of V_K equipped with the IC topology (thus J_K is a topological group, cf. Paragraph 0.1.1). J_K is called **idele group** of K. Elements of J_K are calles **ideles**.

Example I.4.3 Let $K = \mathbb{Q}(i)$. Then we have one Archimedean valuation v_∞ on K (induced by the usual absolute value $|z| = \sqrt{z\overline{z}}$) and one non-Archimedean valuation for any prime $\mathfrak{p} \in \mathbb{Z}[i]$ (compare Theorem 0.5.34 and Theorem 0.5.40). Then $x = (\frac{i}{\pi}, 18, i, 11 - i, 24i, \sqrt{2}, \sqrt{-5}, -5, 13i - 17, \ldots)$ is an idele of $\mathbb{Q}(i)$ (where for all but finitely many $\mathfrak{p}$ we have in fact $x_\mathfrak{p} \in \mathcal{O}_\mathfrak{p}^*$ and we have $x_\mathfrak{p} \neq 0$ for all $\mathfrak{p}$).

$\mathfrak{p}$	∞	$1+i$	3	$2+i$	$2-i$	7	11	$3+2i$	$3-2i$	$\cdots$
$x_\mathfrak{p} \in \mathbb{Q}(i)_\mathfrak{p}$	$\frac{i}{\pi}$	18	i	$11-i$	$24i$	$\sqrt{2}$	$\sqrt{-5}$	-5	$13i-17$	$\cdots$

Table I.4.1: An element of $J_{\mathbb{Q}(i)}$.

The idele group and its topology can be used to deduce Dirichlet's unit theorem (Theorem 0.5.29) and the finiteness of the class number (Theorem 0.5.31), see [Cas67].

If K is a number field, then a continuous homomorphism $\chi : J_K \to \mathbb{C}^*$ with $\chi(K\backslash\{0\}) = 1$ is called a **Hecke character** (compare [RV99, Neu92]). There is an equivalent ideal-theoretic formulation of Hecke characters (without the use of ideles). Sometimes these concepts are being distinguished by calling the ideal-theoretic Hecke characters **Größencharakter**, see [Neu92]. Similarly to the definition of Dirichlet L-functions we can attach an L-function to Hecke characters, the so-called **Hecke *L*-function** $L(s, \chi)$ (for a full definition see [RV99]). If χ is the trivial character χ_0, then $L(s, \chi_0)$ is simply the Dedekind ζ-function of the number field K, i.e., the Riemann ζ-function if $K = \mathbb{Q}$.

In his famous thesis, Tate [Tat67] showed that $L(s, \chi)$ has an analytic continuation and admits a functional equation that relates $L(s, \chi)$ with $L(1 - s, \overline{\chi})$, where $\overline{\chi}$ is the conjugate character of χ, i.e., $\overline{\chi}(x) = \overline{\chi(x)}$.

We should note that these facts were already shown by Hecke in 1920 (compare [Neu92]) without the notion of ideles.

We will see how these results for the Hecke L-function can be used for another L-function in I.12.

I.5

Graph Theory

There is a variety of graphs defined via arithmetic conditions. In this chapter we take a look at some of these graphs and their properties. An important example are unitary Cayley graphs:

Definition I.5.1 Let $n \in \mathbb{N}$. The Cayley graph $X_n := X(\mathbb{Z}/n\mathbb{Z}, (\mathbb{Z}/n\mathbb{Z})^*)$ is called **unitary Cayley graph**.

Here and in the rest of this chapter we regard $\mathbb{Z}/n\mathbb{Z}$ as additive group. Nevertheless, $(\mathbb{Z}/n\mathbb{Z})^*$ denotes the elements in $\mathbb{Z}/n\mathbb{Z}$ that are multiplicatively invertible. Thus a unitary Cayley graph has vertices $\{1, \ldots, n\}$ and two vertices a, b are adjacent if and only if $\gcd(a - b, n) = 1$, i.e., if $a - b$ and n are coprime. In particular, X_n is a complete graph if and only if $n = 1$ or n is prime. Figure I.5.1 shows the unitary Cayley graph X_{12}.

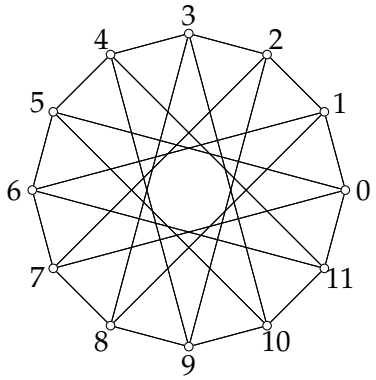

Figure I.5.1: The unitary Cayley graph X_{12}.

Some properties of unitary Caley graphs can be found in [KS07, DG95, Ili09]. These and other properties often depend on arithmetic functions. Some basic facts are

- X_n is regular of degree $\varphi(n)$ and $\kappa(X_n) = \varphi(n)$.
- Let p_* denote the smallest prime dividing n. Then $\chi(X_n) = \omega(X_n) = p_*$.
- The diameter of X_n is

$$\operatorname{diam}(X_n) = \begin{cases} 1, & \text{if } n \text{ is prime} \\ 2, & \text{if } n \text{ is odd, but not a prime} \\ 2, & \text{if } n = 2^\alpha \text{ with } \alpha > 1 \\ 3, & \text{if } n \text{ is even, but not a power of 2} \end{cases}.$$

- The number $T(n)$ of triangles in X_n is

$$T(n) = \frac{n^3}{6} \prod_{p|n} \left(1 - \frac{1}{p}\right)\left(1 - \frac{2}{p}\right).$$

Two of the most important properties of unitary Cayley graphs are the following:

Theorem I.5.2 *Let X_n be a unitary Cayley graph. Then X_n is integral and circulant.*

While the fact that X_n is circulant follows directly by construction, the integrality is a consequence of the integrality of the **Ramanujan sum**

$$c(r,n) = \sum_{\substack{1 \le k \le n \\ \gcd(k,n)=1}} e^{\frac{2\pi i r k}{n}}$$

and the fact that the numbers $c(r,n)(r = 0, \ldots, n-1)$ are exactly the eigenvalues of the adjacency matrix of X_n. More results about the eigenvalues of X_n can be found in [KS07]. In fact the two properties of Theorem I.5.2 hold true for a more general class of graphs, the gcd graphs. Let $\mathcal{D}_n$ denote the set of proper divisors of n, i.e., $\mathcal{D}_n = \{d : 1 \le d \le n-1, d|n\}$.

Definition I.5.3 Let $n \in \mathbb{N}$ and $D \subset \mathcal{D}_n$. For $d \in D$ let

$$G_n(d) := \{dk \in \mathbb{Z}/n\mathbb{Z} : k \in (\mathbb{Z}/n\mathbb{Z})^*\}$$

and $S := \bigcup_{d \in D} G_n(d)$. Then the Cayley graph $X_n(D) := X(\mathbb{Z}/n\mathbb{Z}, S)$ is called **gcd graph**.

Again the vertices are the numbers $\{1, \ldots, n\}$. Here two vertices a, b are adjacent if and only if $\gcd(a - b, n) \in D$. Since $n \notin D$ we obtain a loopless graph. Figure I.5.2 shows the gcd graph $X_{12}(\{3, 4\})$.

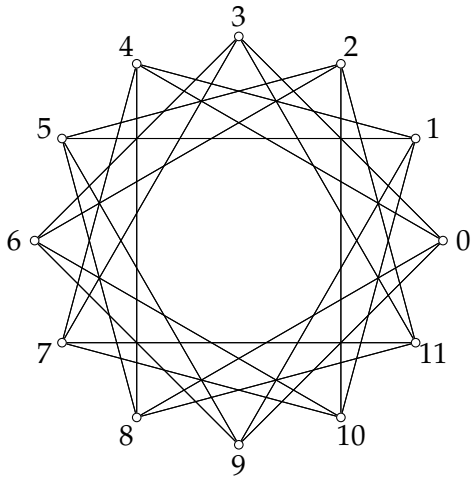

Figure I.5.2: The gcd graph $X_{12}(\{3, 4\})$.

The graph $X_n(D)$ is a complete graph if $D = \mathcal{D}_n$ and we have $X_n = X_n(\{1\})$. If $D = \{d_1, \ldots, d_r\}$, then $X_n(D)$ is connected if and only if $\gcd(n, d_1, \ldots, d_r) = 1$. Note that $(X_n(D))^c = X_n(\mathcal{D} \backslash D)$, thus complements of gcd graphs are again gcd graphs.

So [So06] showed that the two properties in Theorem I.5.2 are in fact characteristic for gcd graphs:

Theorem I.5.4 *A graph is integral and circulant if and only if it is a gcd graph.*

While it is easy to determine some basic characteristics for unitary Cayley graphs (see above), it seems much harder to determine these for general gcd graphs. It is easy to show that $X_n(D)$ is regular of degree $\sum_{d \in D} \varphi(\frac{n}{d})$, but the determination of the diameter, clique number and chromatic number is (at least in general) an open question. At present only special cases and bounds for these values are known (see [SPB12] for the diameter, [BI09] for the clique number and [IB10] for the chromatic number).

Since gcd graphs are defined arithmetically, one should assume that most graph theoretic properties depend on some arithmetic structure. Apart from the basic facts of unitary Cayley graphs that we already mentioned, one more example is the determination of the set $N_2(v)$ for a given vertex v, i.e., the neighbourhood of the neighbourhood of v. This set can be determined through sumsets in

the cyclic group $(\mathbb{Z}/n\mathbb{Z}, +)$, see [SS13]. The characterization generalizes to Cayley graphs of abelian groups. This will be shown in Chapter II.2.

There are of course more graphs defined via arithmetic conditions. One of them is the **coprime graph**:

Definition I.5.5 Let $n \in \mathbb{N}$. The (loopless) coprime graph LCG_n has vertex set $V = \{1, \ldots, n\}$ and two vertices a, b are adjacent if and only if $\gcd(a, b) = 1$ and $(a, b) \neq (1, 1)$.

Some properties of coprime graphs can be found in [ES97, SS09].

In [SK04] the authors investigate an example of an arithmetically defined directed graph H_n. Here there vertex set is $\{0, \ldots, n-1\}$ and there is a directed edge from a to b if $a^2 \equiv b \bmod n$. Figure I.5.3 shows the coprime graph LCG_6 and the graph H_{12}.

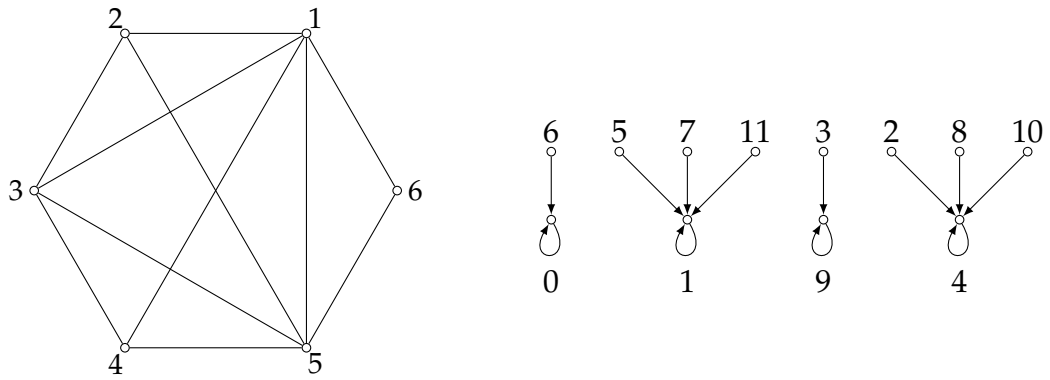

Figure I.5.3: The coprime graph LCG_6 (left) and the graph H_{12} defined in [SK04] (right).

I.6

Linear Algebra

Since linear maps, the main object of study in linear algebra, are in particular additive functions, it seems natural to consider connections between linear algebra and additive number theory. We examine three such connections, namely the Cauchy-Davenport theorem, sums of squares, and the Hopf-Stiefel function.

We start with the Cauchy-Davenport theorem (Theorem 0.5.16) as well as generalizations and applications thereof. The classical proof of Cauchy uses the e-transform of an ordered pair of sets (A, B), which is given by $(A(e), B(e))$ where $A(e) = A \cup (B + e)$ and $B(e) = B \cap (A - e)$ (here G is an abelian group, $A, B \subset G$ and $e \in G$ is an arbitrary element), see [Nat96a].

Apart from the classical proof, the Cauchy-Davenport theorem (as well as generalizations, see, for example, [Pol74]) can be proved with linear algebra, see [DdSH90, CDdS00, Dia04]. This is done with the use of the Kronecker sum:

Theorem I.6.1 *Let V, W be finite dimensional vector spaces over $\mathbb{F}_p$ and let $f : V \to V$ and $g : W \to W$ be linear maps. Let m_φ denote the minimal polynomial of a linear map φ. Then*

$$\deg(m_{f\oplus g}) \geq \min\left\{p, \deg(m_f) + \deg(m_g) - 1\right\}.$$

The original Cauchy-Davenport theorem follows from Theorem I.6.1 by considering diagonal maps, since in this case the degree of m_φ is equal to the number of distinct eigenvalues of φ. Further proofs and some generalizations can be found in [ACGM10].

On the other hand, results similar to the Cauchy-Davenport theorem can be used to deduce results about eigenvalues of the Kronecker product $A \otimes B$ and the Kronecker sum $A \oplus B$ of two complex matrices A, B with distinct eigenvalues, see [Spi82]. Another application of these results is in the theory of linear ordinary differential equations, where one can bound the number of linearly independent solutions of some special differential equations from above, see [Spi82].

Let us now take a look at sums of squares. One of the classical results is the two-squares theorem (Theorem 0.5.10). There is a proof of this theorem based on linear algebra, the famous "one sentence proof", cf. [Zag90]:

Theorem I.6.2 (two-squares theorem for primes) *Let p be a prime. If $p \equiv 1 \bmod 4$, p is representable as a sum of two squares.*

Proof. The involution on the finite set $S = \{(x, y, z) \in \mathbb{N}^3 : x^2 + 4yz = p\}$ defined by

$$(x, y, z) \mapsto \begin{cases} (x + 2z, z, y - x - z), & x < y - z \\ (2y - x, y, x - y + z), & y - z < x < 2y \\ (x - 2y, x - y + z, y), & x > 2y \end{cases}$$

has exactly one fixed point, so $|S|$ is odd and the involution $(x, y, z) \mapsto (x, z, y)$ also has a fixed point. q.e.d.

A variant of this proof can be found in [HB84]. This uses group actions on the set S. For similar proofs regarding representability of primes as $ax^2 + by^2$ see [Els10].

Now we consider sums of squares more conceptually. There are three classical results about sums of squares, namely the two-squares theorem, the three-squares theorem, and the four-squares theorem. Of these theorems, the three-squares theorem has a different structure in its proof. The reason lies in a multiplication formula: If x, y are sums of two, respectively four, squares, then the same holds for their product xy. This is not true for three squares, as the example $x = 3, y = 5$ shows.

The question we could ask is: For which n does such a multiplication formula exist? The answer to this question depends on which kind of multiplication formula we are looking for.

Hurwitz [Hur98] (see also [Coni]) examined multiplication formulae, where the right-hand side consists of bilinear functions, as in the formulae for two and four squares:

Theorem I.6.3 (Hurwitz's theorem) *Let K be a field with* $\mathrm{char}(K) \neq 2$ *and $x_i, y_i \in K$ for $1 \leq i \leq n$. If there are $z_i, 1 \leq i \leq n$, such that every z_i is a bilinear function in $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ and the equation*

$$(x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2) = z_1^2 + \cdots + z_n^2$$

holds, then $n \in \{1, 2, 4, 8\}$.

Not surprisingly, these values are exactly the natural numbers n such that a finite-dimensional real division algebra with dimension n exists, cf. [EHH$^+$92]. In fact, the theorem of Hurwitz can be proved with the structure theorem on division algebras, see [EHH$^+$92].

The original proof of Hurwitz is based on linear transformations. First it is shown with the use of determinants that n must be 1 or even. Further, for dimensional reasons, n cannot be bigger than 8. The only remaining case, $n = 6$, can be examined with the theory of eigenvectors and eigenspaces. For another proof based on vector products see [Coni].

If one does not require the z_i to be bilinear in $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$, there are in fact more values for n such that a multiplication formula exists, cf. [Pfi65]:

Theorem I.6.4 *Let K be a field, $n = 2^m$ and $x_i, y_i \in K$ for $1 \leq i \leq n$. Then there are $z_i, 1 \leq i \leq n$, such that the equation*

$$(x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2) = z_1^2 + \cdots + z_n^2$$

holds.

Here the z_i can even include fractions. For a comparison with the result of Hurwitz, as well as a converse statement and more, see [Cone]. A formula for the case $n = 16$ can be found in [ZE66].

In general one is interested in so-called composition formulae:

Definition I.6.5 Let $r, s, n \in \mathbb{N}$. A **composition** formula of type $[r, s, n]$ is a formula

$$(x_1^2 + \cdots + x_r^2)(y_1^2 + \cdots + y_s^2) = z_1^2 + \cdots + z_n^2$$

where x_i, y_i are variables (mostly over some field K) and z_i are bilinear functions in $(x_1, \ldots, x_r)$ and $(y_1, \ldots, y_s)$.

For more about composition formulae and techniques used to characterize them, see [Sha00b]. Some results can be found in [EHH+92].

We want to get to know a function that gives some results about composition formulae, the **Hopf-Stiefel function**. This will also build a bridge back to the Cauchy-Davenport theorem as well as show an application of additive number theory in linear algebra.

For $r, s \in \mathbb{N}$, the Hopf-Stiefel function $\circ$ is defined by $r \circ s := \beta_2(r, s)$ where for a prime p we let

$$\beta_p(r, s) := \min\{n \in \mathbb{N} : (x + y)^n = 0 \text{ in } \mathbb{F}_p[x, y]/(x^r, y^s)\}.$$

Since each summand of $(x + y)^{r+s-1}$ is divisible by either x^r or y^s, the function $\beta_p(r, s)$ is well-defined and we have $\beta_p(r, s) \leq r + s - 1$. For example, we have $2 \circ 2 = 2$, since $(x + y)^2 = 0$ in $\mathbb{F}_2[x, y]/(x^2, y^2)$. It is clear, that $r \circ s \geq \max\{r, s\}$, but equality does not hold in every case. For example, we have $(x + y)^5 = xy^4$ in $\mathbb{F}_2[x, y]/(x^3, y^5)$, i.e., $3 \circ 5 > 5$. In fact, we have $3 \circ 5 = 7$.

Plagne [Pla03] derived a formula for $\beta_p(r, s)$:

Theorem I.6.6 *For any prime p, we have*

$$\beta_p(r, s) = \min_{t \in \mathbb{N}_0} \left\{ \left(\left\lceil \frac{r}{p^t} \right\rceil + \left\lceil \frac{s}{p^t} \right\rceil - 1 \right) p^t \right\}.$$

Here $\lceil x \rceil$ denotes the ceiling function $\lceil x \rceil := \min\{k \in \mathbb{Z} : x \leq k\}$. Since the minimum is being attained for some t with $0 \leq t \leq \left\lceil \frac{\log(\max\{r,s\})}{\log p} \right\rceil$, this formula is indeed useful. In our examples above, we only have to check $t \in \{0, 1\}$ to get $2 \circ 2 = 2$ and $t \in \{0, 1, 2, 3\}$ to get $3 \circ 5 = 7$.

This formula can be proved by using a result about sumsets in the groups $\mathbb{Z}/n\mathbb{Z}$, which is a generalization of the Cauchy-Davenport theorem. Conversely, when considering an additive problem on subsets of vector spaces V over $\mathbb{F}_p$ for some prime p, the function $\beta_p(r, s)$ plays a role in some Cauchy-Davenport type result (see [Yuz81] for the special case $p = 2$, i.e., the Hopf-Stiefel function, and [EK98] for the general case):

Theorem I.6.7 *Let V be a vector space over $\mathbb{F}_p$ for some prime p and $A, B \subset V$ be nonempty. If $|A| = r$ and $|B| = s$, then $|A + B| \geq \beta_p(r, s)$. Conversely, if V is a vector space over $\mathbb{F}_p$ and $r, s \in \mathbb{N}$ with $\max\{r, s\} \leq |V|$, then there are $A, B \subset V$ with $|A| = r, |B| = s$ and $|A + B| = \beta_p(r, s)$.*

For more information about the Hopf-Stiefel function see [Pla03, EK05] and the references given there. The relation of the Hopf-Stiefel function with composition formulae gets clear with **Pfister's theorem** (sometimes called **Hopf's theorem**), cf. [Pfi65]:

Theorem I.6.8 (Pfister's theorem) *Let $r, s \in \mathbb{N}$ and $K = \mathbb{Q}(x_1, \ldots, x_r, y_1, \ldots, y_s)$, where the x_i, y_j are variables. Then the smallest n such that there is a composition formula of type $[r, s, n]$ in K is $n = r \circ s$.*

For more connections between the Hopf-Stiefel function and composition formulae, see [Sha00b].

In Chapter II.4 we will see another, completely different, connection between number theory and linear algebra. We will consider a puzzle that can be modeled with linear algebra. To solve the puzzle we have to examine the determinants of certain matrices. This can (at least partially) be done with (algebraic) number theory, in particular we need the splitting behaviour of prime ideals in cyclotomic fields.

I.7

Geometry

A lot of number theoretic problems are motivated by geometric problems. Two of the most prominent examples are **Pythagorean triples** (i.e., integer solutions of the equation $x^2 + y^2 = z^2$, cf. Theorem 0.5.4) and **congruent numbers** (i.e., natural numbers n such that the system $n = \frac{1}{2}xy$, $x^2 + y^2 = z^2$ has a rational solution, cf. Figure I.7.1). Although the parametrization of Pythagorean triples given in Theorem 0.5.4 yields congruent numbers, it is in general hard to decide whether a given number n is congruent. We will adress both problems in Chapter I.12, for a first glimpse see [Conf, Conh].

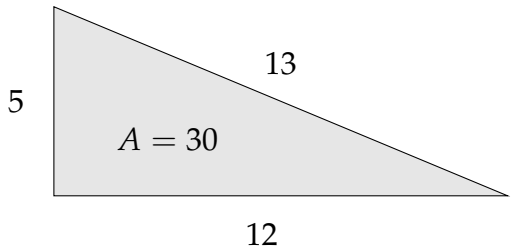

Figure I.7.1: The Pythagorean Triple $(5, 12, 13)$ shows that 30 is a congruent number.

Instead of looking at geometric problems, we will take a look at applications of geometry in number theory. This is called **geometry of numbers** or sometimes **Minkowski theory**. Its starting point is Minkowski's convex body theorem, see [SF07].

Theorem I.7.1 (Minkowski's convex body theorem) *Let Λ be a lattice in $\mathbb{R}^n$ and $S \subset \mathbb{R}^n$ be a convex, centrally symmetric set. If $\mathrm{vol}(S) > 2^n \mathrm{vol}(\Lambda)$, then S contains a nonzero lattice point of Λ.*

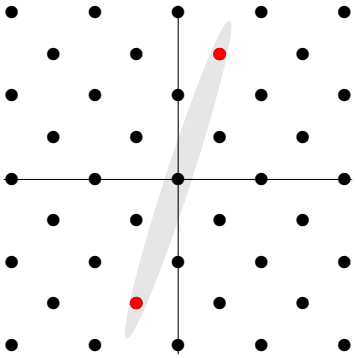

Figure I.7.2: Visualization of Minkowski's convex body theorem.

In most sources the set S is required to be bounded. But since any convex, centrally symmetric and unbounded set S with $\mathrm{vol}(S) > 0$ has in fact volume ∞ and contains convex, centrally symmetric bounded subsets with arbitrarily large volume, the theorem is also true in the unbounded case. Figure I.7.2 shows a visualization of Minkowski's convex body theorem.

There are quite remarkable applications of Minkowski's convex body theorem. We will just mention some important results and indicate how these can be proved with Minkowski's convex body theorem. A more detailled exposition, as well as more examples, can be found in [Cla].

We start with theorems about sums of squares, cf. Paragraph 0.5.3.

Theorem I.7.2 (two-squares theorem for primes) *Every prime p with $p \equiv 1 \bmod 4$ can be written in the form $p = x^2 + y^2$.*

The two-squares theorem for primes can be proved by considering the circle with center at the origin and radius $\sqrt{2p}$ and the lattice

$$\Lambda = \left\{ k_1 \begin{pmatrix} p \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} u \\ 1 \end{pmatrix} : k_i \in \mathbb{Z} \right\}$$

where $u \in \{0, \ldots, p-1\}$ is chosen such that $u^2 \equiv -1 \bmod p$. See Figure I.7.3 for $p = 13$.

Related to the two-squares theorem is the question about the representability of primes in the form $x^2 + ny^2$. Partial results can be obtained with Minkowski's convex body theorem. We will discuss this topic further in Chapter I.8.

Theorem I.7.3 *Let $n \in \mathbb{N}$ and p be an odd prime. If $\left(\frac{-n}{p}\right) = 1$, there are $x, y, k \in \mathbb{Z}$ with $x^2 + ny^2 = kp$ and $1 \leq k \leq \left\lfloor \frac{4\sqrt{n}}{\pi} \right\rfloor$.*

If $n \in \{1, 2\}$, the bound on k implies $k = 1$, so we get a representation of p itself. Here we use almost the same setting as in the proof of the two-squares theorem: We let $u \in \{0, \ldots, p-1\}$ with $u^2 \equiv -n \bmod p$ and consider the ellipse $x^2 + ny^2 = \frac{4\sqrt{n}}{\pi} p$ and the lattice

$$\Lambda = \left\{ k_1 \begin{pmatrix} p \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} u \\ 1 \end{pmatrix} : k_i \in \mathbb{Z} \right\}.$$

Figure I.7.4 illustrates the proof for $p = 17, n = 2$.

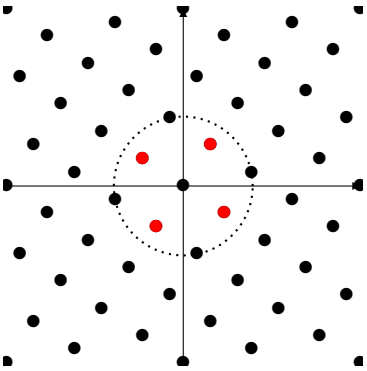

Figure I.7.3: The two-squares theorem for $p = 13$ via Minkowski's convex body theorem.

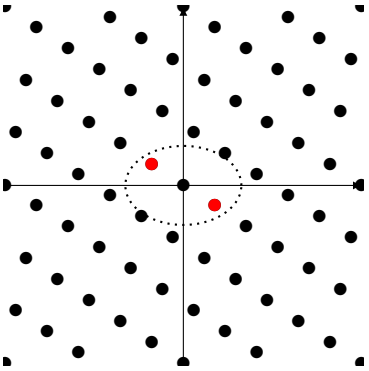

Figure I.7.4: 17 can be written as $x^2 + 2y^2$.

With Minkowski theory, the four-squares theorem for primes can be proved almost analogously to the two-squares theorem.

Theorem I.7.4 (four-squares theorem for primes) *Every prime p can be written in the form $p = w^2 + x^2 + y^2 + z^2$.*

Here we let $r, s \in \{0, \ldots, p-1\}$ with $r^2 + s^2 \equiv -1 \bmod p$ and consider the circle with center at the origin and radius $\sqrt{2p}$ and the lattice

$$\Lambda = \left\{ k_1 \begin{pmatrix} p \\ 0 \\ 0 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} 0 \\ p \\ 0 \\ 0 \end{pmatrix} + k_3 \begin{pmatrix} r \\ s \\ 1 \\ 0 \end{pmatrix} + k_4 \begin{pmatrix} s \\ -r \\ 0 \\ 1 \end{pmatrix} : k_i \in \mathbb{Z} \right\}.$$

As already mentioned (and examined) in Chapter I.6, the structure of the proof of the three-squares theorem is completely different. There are many ways

to prove the three-squares theorem. One of them is to consider **Legendre's equation**, see [SF07]. Again, this can be done with Minkowski's convex body theorem:

Theorem I.7.5 (Legendre's equation) *Let $a_1, a_2, a_3 \in \mathbb{Z}$ be pairwise coprime, square-free, and such that not all a_i have the same sign. Then the equation*

$$a_1 x_1^2 + a_2 x_2^2 + a_3 x_3^2 = 0$$

has a nontrivial integer solution if and only if $-a_i a_j$ is a quadratic residue modulo a_k for each permutation (i, j, k) of $(1, 2, 3)$.

To prove this, we consider the ellipsoid

$$\left\{ |a_1|\, x_1^2 + |a_2|\, x_2^2 + |a_3|\, x_3^2 \leq \sqrt[3]{\frac{6}{\pi}}\, |a_1 a_2 a_3| \right\}$$

and the lattice

$$\Lambda = \left\{ k_1 \begin{pmatrix} 1 \\ * \\ * \end{pmatrix} + k_2 \begin{pmatrix} 0 \\ a_3 \\ * \end{pmatrix} + k_3 \begin{pmatrix} 0 \\ 0 \\ a_1 a_2 \end{pmatrix} \right\}$$

(here the asterisks denote integers that can be determined via the congruences $u_k^2 \equiv -a_i a_j \bmod a_k$, cf. [SF07] for details). Another approach to the three-squares theorem with Minkowski's convex body theorem can be found in [Cla].

For other proofs of the theorems about sums of squares using Minkowski's convex body theorem (i.e., proofs that use different lattices and sets) and partial characterizations of the primes that can be written as $x^2 + 2y^2$ or $x^2 + 3y^2$ see [SF07].

Minkowski's convex body theorem can also be used for algebraic number theory. Let K denote a number field of degree n. Let further $\sigma_1, \ldots, \sigma_r$ be the r real embeddings of K and $\sigma_{r+1}, \ldots, \sigma_{r+s}$ be s pairwise nonconjugate complex embeddings of K. Then we can embed K in $\mathbb{R}^r \times \mathbb{C}^s \cong \mathbb{R}^n$ via $x \mapsto (\sigma_1(x), \ldots, \sigma_{r+s}(x))$. Note that the image of a nonzero ideal $\mathfrak{a} \triangleleft \mathcal{O}_K$ under this map is a lattice in $\mathbb{R}^n$.

There are two main theorems that we can prove when interpreting ideals as lattices: The finiteness of the class number of K (Theorem 0.5.31) and Dirichlet's unit theorem (Theorem 0.5.29). Here we will consider only the first, for the latter see [Cla, Neu92].

The main part in proving the finiteness of the class number is to show that any nonzero ideal contains an element with "small" norm, cf. [Cla, Neu92].

Theorem I.7.6 *Let $\mathfrak{a} \triangleleft \mathcal{O}_K$ be a nonzero ideal. Then there is a nonzero $a \in \mathfrak{a}$ with*

$$\left|N_{\mathbb{Q}}^K(a)\right| \leq \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n} |d(K)|^{\frac{1}{2}} N(\mathfrak{a}).$$

To prove this, we consider the set

$$\left\{(x_1, \ldots, x_r, z_1, \ldots, z_s) \in \mathbb{R}^r \times \mathbb{C}^s : \sum_{i=1}^r |x_i| + 2\sum_{j=1}^s |z_j| \leq \sqrt[n]{2^{n-r}\pi^{-s} n! \sqrt{d(K)} N(\mathfrak{a})}\right\}.$$

Then S is centrally symmetric and convex with volume $2^{r+s}\sqrt{|d(K)|}N(\mathfrak{a})$. Thus, there is a nonzero element $a \in \mathfrak{a}$ with

$$\begin{aligned}\left|N_{\mathbb{Q}}^K(a)\right| = \prod_{i=1}^r |\sigma_i(a)| \prod_{j=r+1}^{r+s} |\sigma_j(a)|^2 &\leq \left(\frac{1}{n}\sum_{i=1}^r |\sigma_i(a)| + \frac{2}{n}\sum_{j=r+1}^{r+s} |\sigma_j(a)|\right)^n \\ &\leq \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n} |d(K)|^{\frac{1}{2}} N(\mathfrak{a}).\end{aligned}$$

Figure I.7.5 illustrates the proof for $\mathfrak{a} = (4, \sqrt{13}) \triangleleft \mathbb{Z}[\sqrt{13}]$.

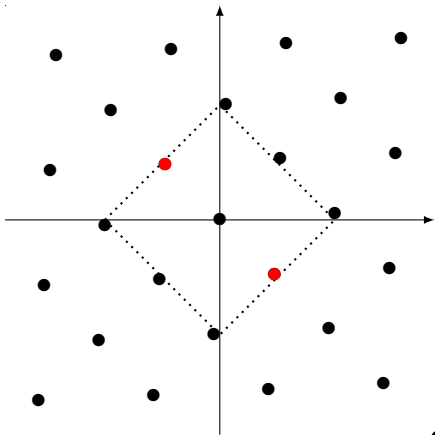

Figure I.7.5: There is an element $a \in (4, \sqrt{13}) \triangleleft \mathbb{Z}[\sqrt{13}]$ with norm less than 7.5953.

There are many more similar results and applications of Minkowski's convex body theorem in number theory, for example results about linear forms, quadratic forms and generalizations to other rings, cf. [Cla].

I.8

Algebra

Similar to analysis, there are a lot of connections between algebra and number theory. We will consider one special aspect that is important for Diophantine equations:

Given a polynomial Diophantine equation $f(x_1, \ldots, x_n) = 0$, we wish to use factorization methods, i.e., we want to factor both sides of the equation and compare the factors. In most cases this is not possible over $\mathbb{Z}$. As an example consider the equation $y^2 = x^3 - 2$, a special case of Mordell's equation. We write this as $x^3 = y^2 + 2$. The left-hand side is already factorized, but the right-hand side cannot be factorized over $\mathbb{Z}$.

The general idea now is to consider ring extensions of $\mathbb{Z}$ and factor the equation over these extensions. In the example above, we can consider $R = \mathbb{Z}[\sqrt{-2}]$ and write $x^3 = (y - \sqrt{-2})(y + \sqrt{-2})$. Since $\mathbb{Z}[\sqrt{-2}]$ is a unique factorization domain (see Theorem 0.5.26) we can look at irreducible factors of both sides. Doing this (compare [IR92]), we see that the equation has only two solutions, namely $(3, 5)$ and $(3, -5)$. In other words, 26 is the only integer such that its predecessor is a square and its successor is a cube. We will examine the equation $y^2 = x^3 + k$ again in Chapter I.12.

Let us take a look at some important equations that can be attacked with this (or a similar) method. More equations (solved with similar or different methods) can be found in [Coh07a].

First we consider another special case of Mordell's equation, namely the equation $y^2 = x^3 \pm 1$. This is related to **Catalan's conjecture**, which implies that the only nontrivial integer solutions of $y^2 = x^3 + 1$ are $(2, \pm 3)$, whereas $y^2 = x^3 - 1$ has no nontrivial integer solutions. More generally, Catalan's conjecture states that 8 and 9 are the only consecutive numbers in the sequence of the powers of natural numbers, i.e., the only nontrivial solutions of the equation $x^n - y^m = 1$ are $(x, y, m, n) = (\pm 3, 2, 3, 2)$. This conjecture has been proved in 2003 by Mihăilescu, cf. [Sch07b]. In his proof he uses the factorization method described above (together with more sophisticated methods). In this case, the equation can be factorized over cyclotomic fields, i.e., fields of the form $\mathbb{Q}(\zeta_p)$ where ζ_p is a p-th root of unity. For more on cyclotomic fields see [Was97].

A more classical (and maybe the easiest interesting) example is the search for Pythagorean triples, i.e., integral solutions of the equation $x^2 + y^2 = z^2$. The characterization given in Theorem 0.5.4 can be proved in many ways (see [IR92, Conf]), one of them is factorizing $x^2 + y^2 = (x + iy)(x - iy)$ in $\mathbb{Z}[i]$.

Generalizing the Pythagorean equation we get to **Fermat's last theorem**, i.e., the consideration of the equation $x^n + y^n = z^n$ with $n \geq 3$. Fermat conjectured that this equation has no nontrivial integer solution. For some special exponents (for example $n = 3$), this can be proved with the factorization method described above, see [HW08, IR92]. More special cases and general approaches can be found in [Edw77, Rib99].

The general equation $x^p + y^p = z^p$ for an odd prime p is much harder. Similar to Catalan's conjecture, we use cyclotomic fields and factor $x^p + y^p$ over $\mathbb{Q}(\zeta_p)$. There are two cases: Either $\gcd(xyz, p) = 1$ or p divides one of x, y, z. The second case is harder to handle but in either case we have the following special case of Fermat's last theroem, which is due to Kummer (see [Edw77]):

Theorem I.8.1 (Fermat's last theorem for regular primes) *If p is regular, then the equation*

$$x^p + y^p = z^p$$

has no nontrivial integer solution.

One of the two key ingredients in the second case is the Kummer-Vandiver conjecture (Conjecture I.4.1). If p is regular, then the Kummer-Vandiver conjecture is true for this prime p and we can use this (together with another consequence of p being regular) to show Fermat's last theorem.

Fermat's last theorem has indeed been proved, but not with cyclotomic methods. The methods used in the proof are described in Chapter I.12.

Let us consider one more equation, namely

$$x^2 \pm ny^2 = c. \tag{I.8.1}$$

We will consider this equation for special values of c. If $c = 1$, this is **Pell's equation**. There are elementary methods to show that the equation $x^2 - dy^2 = 1$ with $d \in \mathbb{Z}$ has infinitely many integer solutions if and only if d is positive and no square, cf. [SF07]. The solutions of Pell's equation can be found with the continued fraction expansion of $\sqrt{d}$.

Pell's equation also has an algebraic interpretation: In the ring $\mathbb{Z}[\sqrt{d}]$, an element $x + \sqrt{d}y$ is a unit if and only if $x^2 - dy^2 = \pm 1$. This is mostly interesting for positive d. Dirichlet's unit theorem (Theorem 0.5.29) tells us that the unit group of the ring of integers of a number field K has rank $r + s - 1$ where r is the number of real embeddings of K in $\overline{\mathbb{Q}}$ and $2s$ is the number of complex embeddings. In the case $K = \mathbb{Q}(\sqrt{d})$ where d is positive and not a square this rank is 1, thus Pell's equation has infinitely many solutions. For a direct proof using only the unit theorem for $\mathbb{Q}(\sqrt{d})$ see [Sch07a]. In fact this method also yields infinitely many solutions of $x^2 - dy^2 = \pm 1$ if the ring of integers of $\mathbb{Q}(\sqrt{d})$ is $\mathbb{Z}[\frac{1+\sqrt{d}}{2}]$.

Another variant of Equation (I.8.1) is the equation $x^2 + ny^2 = p$, where p is a prime. For $n = 1$, this is the two-squares theorem for primes (compare Theorem 0.5.10). For general n we have already seen some partial results in Chapter I.7. Again, this is simple for some values of n but there are a lot of (partially highly complex) different algebraic methods to characterize such primes in general, cf. [Cox13]. These characterizations always involve some congruence condition on p and sometimes conditions about a certain polynomial. The general result is:

Theorem I.8.2 *For $n \in \mathbb{N}$, there is a polynomial $f_n \in \mathbb{Z}[x]$ such that the following holds: Let p be an odd prime that neither divides n nor Δ_{f_n}. Then $p = x^2 + ny^2$ if and only if*

- $\left(\frac{-n}{p}\right) = 1$ *and*
- $f_n(x) \equiv 0 \bmod p$ *has an integer solution.*

Here $\left(\frac{-n}{p}\right)$ denotes the Legendre symbol. The polynomial f_n can be specified, see [Cox13]. The natural numbers n such that the primes of the form $p = x^2 + ny^2$ can be characterized with congruences only (i.e., without needing a polynomial f_n) are called **idoneal** (or **convenient**) numbers, cf. [Kan, Cox13]. There are 65

idoneal numbers known today, and these are assumed to be all. Weinberger [Wei73] showed that there can exist only one more idoneal number, and under assumption of the generalized Riemann hypothesis no other exists. A list of the characterizations of the primes p with $p = x^2 + ny^2$ for $n \leq 29$ can be found in [Jag].

For the cases $n = 1, 2, 3$, the characterization can be proved with the factorization method described above. In these cases, we have

Theorem I.8.3 *Let p be a prime.*

- *p is of the form $p = x^2 + y^2$ if and only if $p = 2$ or $p \equiv 1 \bmod 4$.*
- *p is of the form $p = x^2 + 2y^2$ if and only if $p = 2$ or $p \equiv 1, 3 \bmod 8$.*
- *p is of the form $p = x^2 + 3y^2$ if and only if $p = 3$ or $p \equiv 1 \bmod 3$.*

Interestingly, I was not able to find a reference for such a proof by factorization (unless of course for the case $n = 1$). Therefore, we will discuss the proof in somewhat greater detail. We consider the case $n = 3$ (the other two cases are easier). In fact, we show that for a prime $p > 3$ the following four conditions are equivalent:

1. $p \equiv 1 \bmod 3$,
2. $\left(\frac{-3}{p}\right) = 1$,
3. p is reducible in $\mathbb{Z}[\omega]$ (here $\omega = \frac{-1+\sqrt{-3}}{2}$ is a primitive third root of unity),
4. $p = x^2 + 3y^2$ for suitable $x, y \in \mathbb{Z}$.

The implications $1 \Rightarrow 2$ and $4 \Rightarrow 1$ are easy. If $\left(\frac{-3}{p}\right) = 1$, then p divides $(x - \sqrt{-3})(x + \sqrt{-3})$ in $\mathbb{Z}[\omega]$ for some $x \in \mathbb{Z}$. Since p does not divide one of the factors and $\mathbb{Z}[\omega]$ is a unique factorization domain (see Theorem 0.5.26), p is reducible. If p is reducible, write $p = \alpha\beta$. Then by distinction of cases (depending on the parity of a and b in $\alpha = a + b\omega$) we find a unit $\eta \in \mathbb{Z}[\omega]^*$ such that $\alpha\eta \in \mathbb{Z}[\sqrt{-3}]$. Using the norm in $\mathbb{Z}[\omega]$ then explicitly gives x, y with $x^2 + 3y^2 = p$.

The key steps are the implications $2 \Rightarrow 3$ and $3 \Rightarrow 4$. These work for general n only if (at least) one of the following holds:

- $\mathbb{Z}[\sqrt{-n}]$ is a unique factorization domain.
- The ring of integers R of $\mathbb{Q}(\sqrt{-n})$ is a unique factorization domain and whenever $p = \alpha\beta$ in R, we can choose a unit $\eta \in R^*$ such that $\alpha\eta \in \mathbb{Z}[\sqrt{-n}]$.

The second case cannot occur for $n > 3$ due to Theorem 0.5.26 and the fact that all units in R are roots of unity. On the other hand, $\mathbb{Z}[\sqrt{-n}]$ cannot be a unique factorization domain if $n \geq 3$: Since 2 divides $-n = \sqrt{-n}^2$ for even n and $1+n = (1+\sqrt{-n})(1-\sqrt{-n})$ for odd n, 2 is not a prime in $\mathbb{Z}[\sqrt{-n}]$. On the other hand 2 is irreducible in $\mathbb{Z}[\sqrt{-n}]$ since $x^2 + ny^2 = 2$ is not solvable if $n \geq 3$, thus there are no elements with norm 2 in $\mathbb{Z}[\sqrt{-n}]$. Hence, this method unfortunately works only for $n = 1, 2, 3$.

To conclude this chapter, we note that many number theoretic problems involving congruences can be generalized to problems in factor rings. These can hopefully be solved with algebraic methods and the special case of congruences often has applications in other parts of number theory. We will deal with an example of this approach in Chapter II.3.

I.9

Differential Geometry

In this chapter we discuss connections between number theory and hyperbolic manifolds. In particular, we are interested in special hyperbolic manifolds of dimensions 2 or 3. Since the concepts in dimension 3 require theory that we have not developed, we will just briefly mention one result and then turn our attention to dimension 2.

Of special interest is the volume of manifolds. It is known that there is a hyperbolic 3-manifold of minimal volume, but until today it is neither known which manifold this is, nor what the minimal volume is. For a special class of hyperbolic 3-manifolds, the so-called **arithmetic manifolds**, this problem is solved (we will not describe what an arithmetic manifold is, since this requires too much notation. For an introduction see [MR03]).

In [CFJR01], the authors show that the arithmetic hyperbolic 3-manifold with smallest volume is the **Weeks manifold** (which we will not define here). If α is a root of the polynomial $x^3 - x + 1$ and $K = \mathbb{Q}(\alpha)$, its volume is

$$\frac{3 \cdot \sqrt[3]{23^2}}{4\pi^4} \zeta_K(2) \approx 0.9427073627769.$$

It is conjectured that this is also the smallest volume among all hyperbolic 3-manifolds. In [CFJR01], the authors also determine the next smallest arithmetic hyperbolic 3-manifold. These are the only two having volume less than 1.

Apart from having (presumably) the smallest possible volume among hyperbolic 3-manifolds, arithmetic manifolds are interesting in their own right since their geometric structure is connected with arithmetic information (for example with the Hilbert symbol). For a more detailed exposition, see [MR03].

In dimension 2, there is one special manifold we are interested in, namely one of the standard models of hyperbolic geometry already described in Paragraph 0.3.2: The upper half plane $\mathbb{H}$. We take a look at two concepts. First we describe a tessellation of the upper half plane. After that, we will discuss special functions that obey a certain transformation law under some isometries of $\mathbb{H}$.

We begin with the Ford circles described in Chapter I.2. First, we extend them periodically from $[0,1]$ to all of $\mathbb{R}$. We also add the line $y = 1$, which we will view as the Ford circle $C[1/0]$. Now given any two Ford circles that are tangent, we construct the unique geodesic that intersects the touching boundary points orthogonally. A partial result of this process (for the Ford circles $C[a/b]$ up to $b = 4$) is depicted in Figure I.9.1.

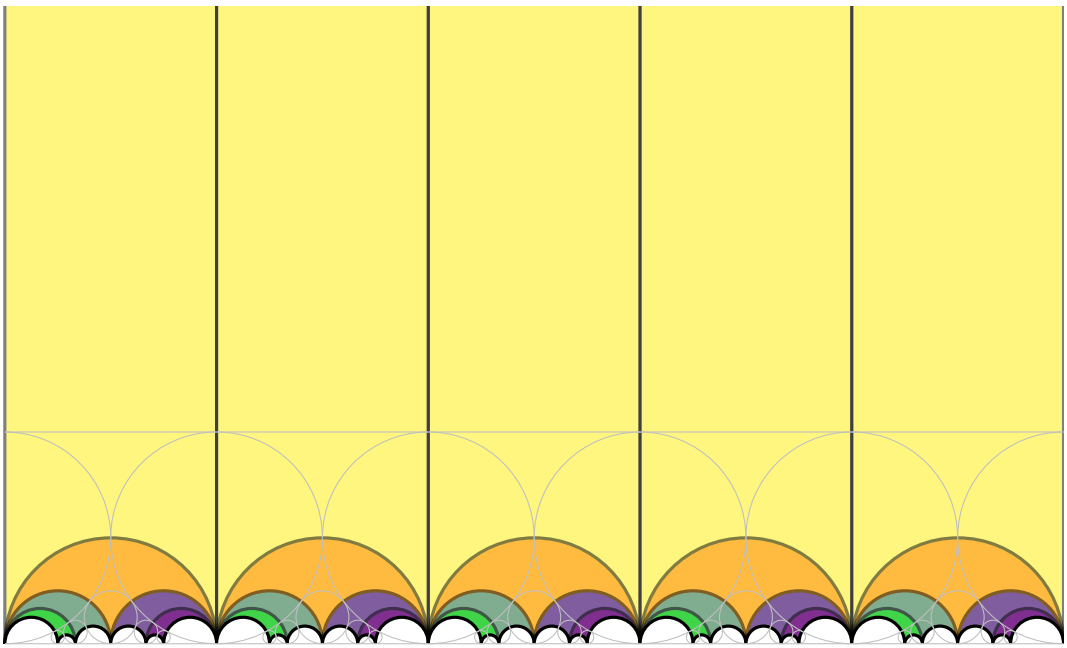

Figure I.9.1: A tessellation of $\mathbb{H}$ via Ford Circles.

Given three Ford circles that are pairwise tangent, these geodesics define a (hyperbolic) triangle whose corners are ideal points. These triangles form a **tessellation** of $\mathbb{H}$, i.e., these triangles cover all of $\mathbb{H}$ and do not overlap. Moreover, this tessellation is invariant under the action of $\mathrm{SL}_2(\mathbb{Z})$ (which is a special subgroup of the group of isometries of $\mathbb{H}$ and will become important soon) on

$\mathbb{H}$. For more about this tessellation and connections to continued fractions see [Ser15, Hat]. Another tessellation of $\mathbb{H}$ that uses Ford circles can be found in [CF97].

We determine a fundamental domain of $\mathbb{H}$ for the action of $\mathrm{SL}_2(\mathbb{Z})$ with help of the above tessellation. Consider the triangle $\triangle$ whose corners are the ideal points $0, 1, \infty$. Note that there is a nontrivial element in $\mathrm{SL}_2(\mathbb{Z})$, namely $S = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$, which maps $\triangle$ to itself, thus $\triangle$ cannot be a fundamental domain. Since S has order 3, we need to divide $\triangle$ in (at least) three subsets to get a fundamental domain (in fact three subsets will suffice, cf. [Ser15]). Since the fixed point of S is $\frac{1+\sqrt{-3}}{2}$, a fundamental domain is given by the set $\mathfrak{E}$ shown in Figure I.9.2. Since any of the boundary lines of $\mathfrak{E}$ is a part of a geodesic, $\mathfrak{E}$ is a quadrangle.

For most applications, this is not the fundamental domain we want to work with. We would like to work with a triangular fundamental domain rather than with a quadrangular. To achieve that, we note that $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ maps each element $z \in \mathbb{H}$ to $z + 1$. Therefore, we can shift a part of $\mathfrak{E}$ by some integer n to the right or left. Shifting the right half of $\mathfrak{E}$ by 1 to the left, we get the fundamental domain $\mathfrak{F}$ that is most often used, see Figure I.9.2.

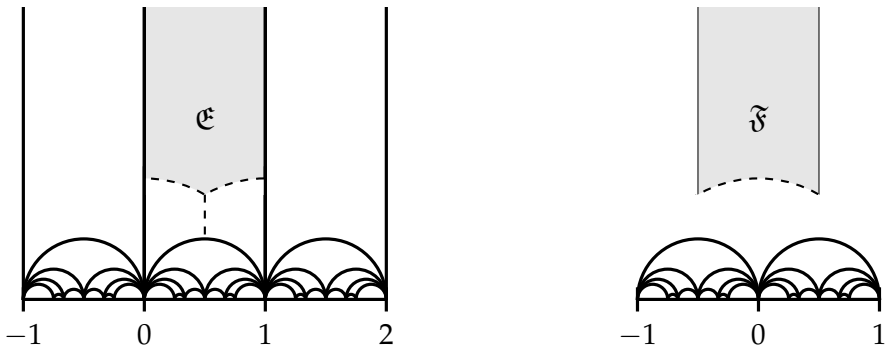

Figure I.9.2: The fundamental domains $\mathfrak{E}$ (left) and $\mathfrak{F}$ (right).

Now we consider functions that are (almost) invariant under $\mathrm{SL}_2(\mathbb{Z})$. Unless stated otherwise, proofs of the results can be found in [KK07].

Definition I.9.1 Let $f : \mathbb{H} \to \mathbb{C}$ be a (not necessarily holomorphic) function and $M \in \mathrm{SL}_2(\mathbb{Z})$. For even $k \in \mathbb{Z}$ let $f|_k M : \mathbb{H} \to \mathbb{C}$ be the function with

$$(f|_k M)(\tau) := \left(\frac{\mathrm{d}M\tau}{\mathrm{d}\tau}\right)^{\frac{k}{2}} f(M\tau).$$

We call f a **weakly modular form of weight k** if f is meromorphic on $\mathbb{H}$ and $f|_k M = f$ for all $M \in \mathrm{SL}_2(\mathbb{Z})$. In this context the group $\mathrm{SL}_2(\mathbb{Z})$ is called **modular group**.

For $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ we have $\frac{\mathrm{d}M\tau}{\mathrm{d}\tau} = (c\tau + d)^{-2}$, thus the condition $f|_k M = f$ can also be written as

$$f(M\tau) = (c\tau + d)^k f(\tau). \tag{I.9.1}$$

This can be further simplified. Let $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Then the modular group $\mathrm{SL}_2(\mathbb{Z})$ is generated by J and T. Thus a meromorphic function $f : \mathbb{H} \to \mathbb{C}$ is a weakly modular form of weight k if and only if

$$f(\tau + 1) = f(\tau) \text{ and } f\left(-\frac{1}{\tau}\right) = \tau^k f(\tau). \tag{I.9.2}$$

This can also be defined for odd k, but due to the transformation rule (I.9.1) the only weakly modular form of odd weight is the zero function.

Due to (I.9.2) any weakly modular form f is periodic modulo 1, therefore we can write f as a Fourier series

$$f(\tau) = \sum_{m \in \mathbb{Z}} a_f(m) e^{2\pi i m \tau}.$$

We say that f is **meromorphic at ∞** if there is an $m_0 \in \mathbb{Z}$ with $a_f(m) = 0$ for all $m \leq m_0$. If we can choose $m_0 = 0$, we call f **holomorphic at ∞**. A weakly modular form which is holomorphic in $\mathbb{H}$ and at ∞ is called **modular form**. If f is a modular form and $f(\infty) = 0$ (i.e., $a_f(0) = 0$), we call f a **cusp form**.

Modular forms (or, to be more precise, their transformation rule) can be motivated even more with differential geometry, cf. [Mil97]. For an ad hoc explanation that this definition is "nice", define $\widetilde{f}(\tau) := \Im(\tau)^{\frac{k}{2}} |f(\tau)|$. Recall that the hyperbolic metric scales with the inverse of the square of the imaginary part, thus this definition connects the function f with hyperbolic information. By a simple calculation we see that $\widetilde{f}$ is invariant under $\mathrm{SL}_2(\mathbb{Z})$. Further, we can introduce a scalar product, the **Petersen scalar product**, on the space of cusp forms that is invariant under $\mathrm{SL}_2(\mathbb{Z})$.

In the theory of modular forms one often uses the fundamental domain $\mathfrak{F}$ that we constructed above. It occurs in many formulae (we will see one of them

below) and some properties of modular forms (such as boundedness or zeros) can be deduced for all of $\mathbb{H}$ if they are known for $\mathfrak{F}$.

Modular forms are an important object in many number theoretic topics. For example, they can be used to deduce formulae for the number of representations of a natural number n as the sum of k squares, see [LR11]. We mention some basic examples and results. More information can be found in [KK07, Mil97].

The set $\mathscr{M}_k$ of all modular forms of weight k and the set $\mathscr{S}_k$ of all cusp forms of weight k are $\mathbb{C}$-vector spaces (with the canonical operations). Their dimensions are given by the **dimension formula**

$$\dim \mathscr{M}_k = \begin{cases} \left\lfloor \frac{k}{12} \right\rfloor, & k \equiv 2 \bmod 12 \\ \left\lfloor \frac{k}{12} + 1 \right\rfloor, & k \not\equiv 2 \bmod 12 \end{cases}, \qquad \dim \mathscr{S}_k = \dim \mathscr{M}_k - 1.$$

If f is a nonzero modular form of weight k, we have the **weight formula**

$$\sum_{w \in \mathfrak{F}} \frac{\operatorname{ord}_w(f)}{\operatorname{ord} w} = \frac{k}{12}.$$

Here $\mathfrak{F}$ is the fundamental domain we constructed above and

$$\operatorname{ord} w := \begin{cases} 2, & w = i \\ 3, & w = \frac{1}{2} + \frac{\sqrt{-3}}{2} \\ 1, & w \neq i, \frac{1}{2} + \frac{\sqrt{-3}}{2} \end{cases}.$$

There are many important examples of (weakly) modular forms, such as Eisenstein series or the discriminant $\Delta(\tau)$. We will just briefly look at the ***j*-invariant**. This is a weakly modular form of weight 0, which can be defined as a quotient of an Eisenstein series and the discriminant. Its Fourier expansion is

$$j(\tau) = e^{-2\pi i \tau} \sum_{m=0}^{\infty} j_m e^{2\pi i m \tau}$$

where the j_m have interesting arithmetic properties (see [KK07, Apo90] for a table of values and some more facts). These values will become important in Chapter I.11.

The j-invariant is not the only modular form whose Fourier coefficients are important. In general, the behaviour of these coefficients, and in particular their growth, has important applications. There are two basic results: For cusp forms we have $\alpha_f(m) = \mathcal{O}(m^{\frac{k}{2}})$, for general modular forms we get $\alpha_f(m) = \mathcal{O}(m^{k-1})$.

More bounds can be found in [Iwa97]. An example for better bounds for special classes of modular forms, as well as an application to expressing natural numbers as sums of three squarefull natural numbers can be found in [Blo04].

When taking subgroups U of $\mathrm{SL}_2(\mathbb{Z})$, we can consider functions such that the transformation rule (I.9.2) holds for elements of U but do not need to hold for other elements of $\mathrm{SL}_2(\mathbb{Z})$. An important example of such subgroups are **congruence subgroups**. A special case of those is the following.

Definition I.9.2 Let $N \in \mathbb{N}$. Then we define

$$\Gamma_0(N) := \left\{ M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : c \equiv 0 \bmod N \right\}.$$

The special case $N = 1$, i.e., $\Gamma_0(1) = \mathrm{SL}_2(\mathbb{Z})$, is also denoted by Γ.

We call a function $f : \mathbb{H} \to \mathbb{C}$ a **modular form of level N** if all of the above axioms for modular forms hold, but the transformation rule has only to be true for $M \in \Gamma_0(N)$. Let $\mathscr{M}_k(N)$ be the space of modular forms of weight k and level N and $\mathscr{S}_k(N)$ be the space of cusp forms of weight k and level N.

Cusp forms of weight 2 and level N will be important in Chapter I.12, together with the so-called Hecke operators. These are linear operators on the space of modular forms. We will just give the definition for primes.

Definition I.9.3 For a prime p and a modular form f of weight k (and level $N \in \mathbb{N}$) the **Hecke operator** T_p is defined by

$$(T_p f)(\tau) := p^{k-1} f(p\tau) + \frac{1}{p} \sum_{b=1}^{p} f\left(\frac{\tau + b}{p}\right).$$

If (for a fixed prime p) a modular form f satisfies $T_p(f) = \lambda_p(f) \cdot f$ for some $\lambda_p(f) \in \mathbb{C}$, the modular form f is called **eigenform** of T_p with **eigenvalue** $\lambda_p(f)$.

For $k \in \{0, 4, 6, 8, 10, 14\}$, the dimension formula yields that the spaces $\mathscr{M}_k$ and $\mathscr{S}_{k+12}$ are one-dimensional. Since the Hecke operators are endomorphisms both of $\mathscr{M}_k$ and $\mathscr{S}_k$ (see [KK07]), the modular forms in theses spaces are eigenforms.

Finally, we mention that we can associate an L-function to modular forms. If f is a modular form with Fourier expansion $\sum_{m=0}^{\infty} a_m(f) e^{2\pi i m\tau}$, we define the L-function attached to f by $L(f,s) := \sum_{m=1}^{\infty} a_m(f) m^{-s}$. Here bounds on the Fourier coefficients are useful, since these determine the convergence of $L(f,s)$. There

are deep connections between the L-function $L(f,s)$ and the Hecke operators T_p. Properties of f with respect to Hecke operators (for example, being an eigenform) will result in properties of $L(f,s)$ (for example, having an Euler product), cf. [DS05, LR11]. This L-function will recur in Chapter I.12.

Conversely, Dirichlet series $\sum_{m=1}^{\infty} a_m m^{-s}$ define (under certain assumptions) a modular form f with Fourier expansion $\sum_{m=0}^{\infty} a_m(f) e^{2\pi i m\tau}$, see [Hus04].

I.10

Complex Analysis

We have already seen (and will see again) one of the main uses of complex analysis in number theory, namely **ζ- and L-functions**. Since these appear in other chapters too, we will just briefly adress this topic. Afterwards, we will turn our attention to another connection.

The theory of ζ- and L-functions follows one general concept: Given an (algebraic, geometric, or arithmetic) object X, we associate to this object a function depending on one complex variable. In some cases, this is done by considering a sequence $(a_n)_{n\in\mathbb{N}}$ of (mostly) integers, that encodes information about the object X, and define the ζ- or L-function to be the Dirichlet series $\sum_{n=1}^{\infty} a_n n^{-s}$. Table I.10.1 shows the ζ- and L-functions that we have considered so far and also the ones we will see in later chapters.

appearance in Chapter	object	function
0, I.2, A.1	$\mathbb{Q}$ resp. the sequence $a_n = 1$	Riemann ζ-function $\zeta(s)$
0, A.1	Dirichlet character χ	Dirichlet L-function $L(s,\chi)$
0, I.9	number field K	Dedekind ζ-function $\zeta_K(s)$
I.1, I.11	the field $\mathbb{F}_q(T)$	Carlitz ζ-function $\zeta_p(s)$
I.4	projective algebraic variety X	$\zeta(X,s)$
I.4, I.12	Hecke character χ	Hecke L-function $L(s,\chi)$
I.9, I.12	modular form f	$L(s,f)$
I.12	elliptic curve E	Hasse-Weil L-function $L(E,s)$

Table I.10.1: List of ζ- and L-functions addressed in this thesis.

For many more different ζ- and L-functions, see [Wat, LMF, Coh07b].

If a function is a Dirichlet series, i.e., it is of the form $\sum_{n=1}^{\infty} a_n n^{-s}$ for some complex sequence $(a_n)_{n\in\mathbb{N}}$, it will converge in some right half plane $\sigma \geq \sigma_0$ where $\sigma = \Re(s)$ and $\sigma_0 \in \mathbb{R}$. Further, it has an Euler product if $(a_n)_{n\in\mathbb{N}}$ is multiplicative, cf. [Brü95].

In most cases, one wishes for the ζ- or L-function to have some nice properties, for example analytic or meromorphic continuation to all of $\mathbb{C}$ and a functional equation, see also Chapter I.4.

ζ- and L-functions and their values have important applications in number theory. One of the first is the use of the Riemann ζ-function in the prime number theorem, see [Brü95]. For some applications of special values of ζ-functions, see [Zag94].

One could ask in which cases we call the attached function a ζ-function and in which cases we refer to an L-function. It seems that there is no particular reason for the one or the other choice, see [Cor].

The topic which we will mainly adress here is the **circle method**, also called **Hardy-Littlewood method**. This method can be used to get asymptotic formulae for the number of solutions of Diophantine equations. Before describing some of the problems that can be attacked, we present the main idea behind the circle method, cf. [Vau97]. See Appendix A.2 for an introductory example.

The historically first version of the circle method is the following: Let $A \subset \mathbb{N}$ (where A is assumed to be infinite) and denote by $(a_m)_{m\in\mathbb{N}}$ the strictly increasing sequence of all elements of A. We want to know in how many ways a natural number n can be written as a sum of s elements of A (considering the order of selection). Let $R_s(n)$ denote this number.

To a_m, we associate the series

$$F(z) := \sum_{m=1}^{\infty} z^{a_m}$$

where $|z| < 1$. The s-th power of $F(z)$ is exactly

$$F(z)^s = \sum_{n=1}^{\infty} R_s(n) z^n.$$

To compute $R_s(n)$ (at least asymptotically and for big n), we use the Cauchy's integral formula (Theorem 0.2.8) to get

$$R_s(n) = \frac{1}{2\pi i} \int_{|z|=\rho} \frac{F(z)^s}{z^{n+1}} \, \mathrm{d}z$$

for any $\rho \in (0,1)$. This is the reason for the name circle method.

Nowadays, one often uses a refined version. To ease notation, we will write $e(\alpha) = e^{2\pi i \alpha}$ in the rest of this chapter, as usual in analytic number theory. Now we consider the Diophantine equation

$$D(x_1, \ldots, x_k) = n, \text{ where } D(x_1, \ldots, x_k) = a_1 x_1^{m_1} + \cdots + a_k x_k^{m_k}. \tag{I.10.1}$$

For simplicity we assume that $a_i \in \mathbb{N}$ and $2|m_i$ for all i. Then each summand in $D(x_1, \ldots, x_k)$ is positive and thus x_i is bounded for every i. This is not necessary for the circle method to work, cf. Appendix A.2.

We are interested in the number $R_N(n)$ of solutions of the above equation with $|x_i| \leq N$ for all i (in most cases, the exponents m_i in the polynomial D are all the same. Then we choose $N = n^{\frac{1}{m_i}}$). Later, we consider the limit $N \to \infty$ and define $R(n) = \lim_{N\to\infty} R_N(n)$. Let $f_i(\alpha) = \sum_{|l_i|\leq N} e(a_i l_i^{m_i} \alpha)$. If N is chosen large enough (such that each x_i has to be smaller than N to satisfy Equation (I.10.1)), we get

$$\prod_{i=1}^{k} f_i(\alpha) = \sum_{i=1}^{k} \sum_{|l_i|\leq N} e(\alpha \cdot D(x_1, \ldots, x_m)) = \sum_{m=1}^{M} e(\alpha m) R(m)$$

for some M. With the **orthogonality relation**

$$\int_0^1 e(\alpha m) \, \mathrm{d}\alpha = \begin{cases} 1, & m = 0 \\ 0, & m \in \mathbb{Z} \backslash \{0\} \end{cases}$$

we get

$$\begin{aligned} R(n) &= \int_0^1 \sum_{m=1}^{M} R(m) e(\alpha m) e(-\alpha n) \, \mathrm{d}\alpha \\ &= \int_0^1 \sum_{m=1}^{M} e(\alpha (D(x_1, \ldots, x_k) - n)) \, \mathrm{d}\alpha \\ &= \int_0^1 \prod_{i=1}^{k} f_i(\alpha) e(-\alpha n) \, \mathrm{d}\alpha. \end{aligned}$$

For other applications (such as sums of primes) the functions f_i can be varied accordingly. Now $f_i(\alpha)$ is "big" if α is close to some member of a Farey sequence,

i.e., if α is close to $\frac{a}{q}$ for some "small" q (the optimal value of q, depending on N, has to be determined in each individual use of the circle method). We split the integral according to the choice of q into two parts: the **major arcs**, i.e., values α close to some fraction $\frac{a}{q}$ with small q and the **minor arcs**, i.e., all the other α.

If q is chosen well, then the minor arcs contribute (asymptotically) less than the major arcs, so their contribution goes into some error term. On the major arcs we can try to expand f asymptotically in a finite series.

For another version of the circle method see [HB96]. In each case, the aim is to achieve a formula of the form

$$R(n) = \mathfrak{S}(n)J(n) + \mathcal{O}(n^t),$$

where $\mathfrak{S}(n)$ is an infinite series involving exponential sums and $J(n)$ is an integral. We call $\mathfrak{S}(n)$ and $J(n)$ the **singular series** and **singular integral**, respectively. The singular series encodes the arithmetic properties of the equation, for example $\mathfrak{S}(n) = 0$ if and only if the equation considered is not solvable. The singular integral encodes the analytic properties of the equation, for example the number of solutions if the equation is solvable.

The main part of the circle method is to show the following: The singular series vanishes if and only if the equation considered is unsolvable. In the other case, this series converges to a nonnegative value. Together with the singular integral this gives an asymptotic formula if the order of $J(n)$ is greater than t.

The singular integral is usually estimated with integral or summation formulae such as the Euler-Maclaurin formula. The singular series consists of summands that, in some cases, are multiplicative. The techniques here involve general techniques about exponential sums as described, for example, in [GK91]. Sometimes, one also associates a ζ-function to the summands of $\mathfrak{S}(n)$, for which we can use Perron's formula and the residue theorem.

The most prominent applications of the circle method are Waring's problem (Problem 0.5.13, respectively the variant mentioned) and the ternary Goldbach problem (Problem 0.5.14), see [Vau97]. But there are many other additive problems that can be handled with the various forms of the circle method, such as representation numbers of quadratic forms (and with it, the Hasse-Minkowski theorem (Theorem I.1.4)), see [HB96].

We conclude this chapter with a note about the Farey sequence. This sequence has already been mentioned above, when we distinguished between major and

minor arcs. Of course, one does not need the Farey sequence itself for this distinction, so we could have worked without knowing what a Farey sequence is. But there is indeed a use of Farey sequences in a variant of the circle method. To be more precise, this uses not the Farey sequence, but the Ford circles we associated to it. In fact, one can use a part of their boundaries as an integration contour for results about the partition function, see [Rad43]. Later, this has been refined to get exact formulae for Fourier coefficients of modular forms, see [Apo90].

I.11

Physics and Computer Science

Now we show some connections between number theory and two fields close to mathematics, namely physics and computer science.

In Chapter I.2 we have already seen some applications of number theory in physics. There are some more connections, but since most of them would require too much further understanding (in particular in physics), we will just mention some concepts without explaining them.

One connection we want to mention is **moonshine**. For a comprehensive introduction see [Gan06b]. Moonshine connects physics, representation theory and modular forms. First, the connection between representation theory and modular forms has been noticed by the fact that the Fourier coefficents of $j(\tau)$ (compare Chapter I.9) are linear combinations of the dimensions of the irreducible representations of the **monster group** (the biggest sporadic finite simple group with order approximately $8 \cdot 10^{53}$), see [CN79]. For existence and properties of the Monster group see [Gri82]. It has later been shown that similar statements hold for other simple groups, see [Gan06a].

Later, possible connections between these concepts and physics have been discovered (see [DGO15] for a connection to quantum gravity and some open problems, and [Gan06a] for a connection to conformal field theory). For more references, see [Gan06a, DGO15].

There are more connections between number theory and physics. See [HJM15] for a reformulation of the Riemann hypothesis with string theory and [SH11] for another connection between physics and the Riemann hypothesis. A collection of different connections between number theory and physics can be found in [Mar].

Concerning the connections between number theory and computer science we can be much more specific. It seems natural that there are connections, since whenever we have a natural number n, we write this number in some base b (mostly 10). In this base, a natural number is just a string, i.e., an element in Σ^* for $\Sigma = \{0, \dots, b-1\}$. Hence we will always have alphabets that consist of integers. In this case we refer to the integers in the alphabet as **digits**. We consider two connections. One of the two connections will be shown here, another in Chapter II.1. The presentation here follows [Sha04].

To begin, we need the notion of automatic sequences. Let $b \geq 2$ be a natural number, $\Sigma = \{0, \dots, b-1\}$ and $M = (Q, \Sigma, \delta, q_0, \Delta, \eta)$ some DFAO. For any $n \in \mathbb{N}$ let $\langle n \rangle_b$ be the representation of n in base b and compute $a_n = \eta(\delta(q_0, \langle n \rangle_b)) \in \Delta$ (it is not important if we start with the first or the last digit of n, as long as we proceed the same way for every n, see [Fog02]). This yields a sequence $(a_n)_{n\in\mathbb{N}}$. Any sequence that can be constructed in this way is called ***b*-automatic**. If a sequence $(a_n)_{n\in\mathbb{N}}$ is b-automatic and M is a DFAO that yields $(a_n)_{n\in\mathbb{N}}$, then we say that M is a ***b*-automaton** for $(a_n)_{n\in\mathbb{N}}$.

One of the most important theorems connecting formal languages and number theory is **Christol's theorem**, see [Chr79]:

Theorem I.11.1 (Christol's theorem) *Let q be a prime power and $(u_n)_{n\in\mathbb{N}}$ be a sequence over $\mathbb{F}_q$. Then the formal power series $\sum_{n=1}^{\infty} u_n X^n$ is algebraic over $\mathbb{F}_q[X]$ if and only if $(u_n)_{n\in\mathbb{N}}$ is q-automatic.*

Example I.11.2 Let p be a prime and $q = p^s$. Consider the DFAO

$$M = (\{0, \dots, q-1\}, \{0, \dots, q-1\}, \delta, 0, \{0, \dots, p-1\}, \eta)$$

with $\delta(s, \sigma) = \sigma$ and $\eta(s) \equiv s \bmod p$ (see Figure I.11.1 for $p = q = 3$). If n has base q representation $\langle n \rangle_q = b_k \dots b_0$, we have

$$a_n = \eta(\delta(0, \langle n \rangle_q)) = \eta(\delta(b_1, b_0)) = \eta(b_0) \equiv b_0 \bmod p,$$

i.e., $(a_n)_{n\in\mathbb{N}}$ is a sequence with $a_n \equiv n \bmod p$. Thus for all q the power series $\sum_{n=0}^{\infty} nX^n$ is algebraic over $\mathbb{F}_q[X]$.

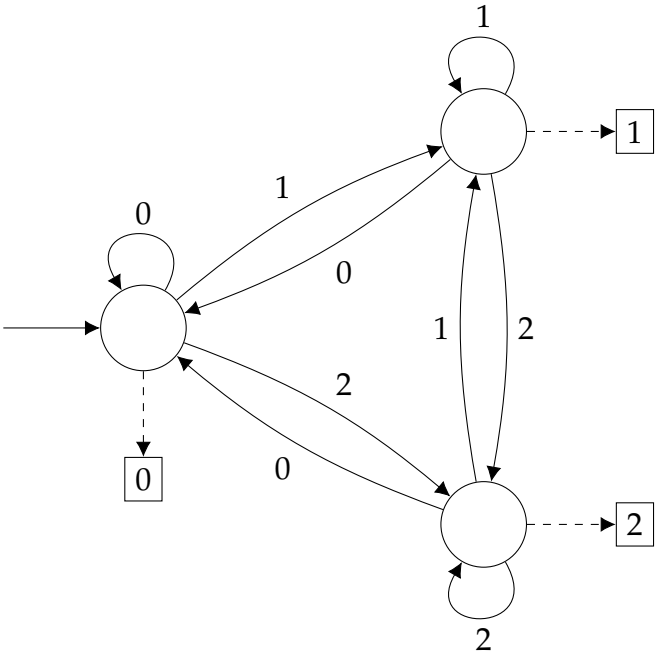

Figure I.11.1: A DFAO for the automatic sequence $a_n \equiv n \bmod 3$.

Before considering applications, we make a general note about algebraic power series and Laurent series. Let $k \in \mathbb{N}$ and $(a_n)_{n \geq -k}$ be a sequence in some field K. Consider the four series

$$f_1 = \sum_{n=0}^{\infty} a_n X^n, \quad f_2 = \sum_{n=-k}^{\infty} a_n X^n, \quad f_3 = \sum_{n=0}^{\infty} a_n X^{-n}, \quad f_4 = \sum_{n=-k}^{\infty} a_n X^{-n},$$

i.e., $f_1 \in K[[X]], f_2 \in K((X)), f_3 \in K[[X^{-1}]], f_4 \in K((X^{-1}))$. Then if one of these series is algebraic over $K[X]$, any of the other three is algebraic, too. It is trivial that f_1 is algebraic if and only if f_2 is algebraic, since these two series differ only by a summand $h \in K[X]$. The same holds for f_3 and f_4. In fact, these equivalences also follow from Christol's theorem (at least for $K = \mathbb{F}_q$), since a sequence is b-automatic if it differs only on finitely many places from a b-automatic sequence (this can be seen by taking a DFAO for the sequence that is known to be b-automatic, appending states for any of the changed places, and varying the transition function accordingly). It is also trivial that f_1 is algebraic if and only if f_3 is algebraic, since every polynomial $p \in (K[X])[t]$ with $p(f_1) = 0$ also satisfies $p(f_3) = 0$.

With this observation, we examine both power series and Laurent series, either in X or X^{-1}, with Christol's theorem, and any statement that we make for one of the respective cases also holds for the others.

With Christol's theorem we can deduce properties of derived power series. Assume f and g are both algebraic over $\mathbb{F}_q[X]$. Then their Hadamard product $f \odot g$ is also algebraic, see [Tha] (this is in general not true in characteristic 0). To prove this, let

$$f = \sum_{n=0}^{\infty} f_n X^n, \quad g = \sum_{n=0}^{\infty} g_n X^n$$

and

$$M_f = (Q_f, \Sigma, \delta_f, q_{0,f}, \Delta, \eta_f) \text{ and } M_g = (Q_g, \Sigma, \delta_g, q_{0,g}, \Delta, \eta_g)$$

be the q-automata for $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$, respectively. Consider the DFAO

$$M = (Q_f \times Q_g, \Sigma, \delta, (q_{0,f}, q_{0,g}), \Delta, \eta)$$

with

$$\delta((q_f, q_g), \sigma) := (\delta_f(q_f, \sigma), \delta_g(q_g, \sigma)) \text{ and } \eta((q_f, q_g)) := \eta_f(q_f) \cdot \eta_g(q_g).$$

Then M is a q-automaton for $(a_n)_{n\in\mathbb{N}} := (f_n \cdot g_n)_{n\in\mathbb{N}}$, thus $f \odot g$ is algebraic over $\mathbb{F}_q[X]$.

If f is algebraic over $\mathbb{F}_q[X]$, the same holds for its formal derivative, since for $f = \sum_{n=0}^{\infty} f_n X^n$ we have

$$f' = X^{-1} \sum_{n=0}^{\infty} n f_n X^n = X^{-1} f \odot g$$

with $g = \sum_{n=0}^{\infty} n X^n$. Since we have already shown above that g is algebraic, and we know that the Hadamard product of two algebraic power series is again algebraic, f' is algebraic.

Until now we have shown that certain power series are algebraic. Now we show the transcendence of a power series $f = \sum_{n=0}^{\infty} f_n X^n$. This is in general harder, since we need to show that no polynomial p with $p(f) = 0$ exists. Even with Christol's theorem this seems to remain true: To show that a power series is algebraic, we need to construct a DFAO M that is a q-automaton for $(f_n)_{n\in\mathbb{N}}$. To show that a power series is transcendental, we have to show that no such DFAO can exist. In some cases this is not too hard, since k-automatic sequences have some special properties.

For this, let $\Sigma = \{0, \ldots, q-1\}$ and $(a_n)_{n\in\mathbb{N}} \subset \Delta$ be q-automatic for some finite set Δ. Then the subsequence $(a_{q^n-1})_{n\in\mathbb{N}}$ is eventually periodic. An elaboration of this can be found in [Lot05]. We can also confirm this with a simple observation.

If $c(n)$ is any sequence such that the base q representation of $c(n)$ is n times a fixed digit $d \in \{0, \ldots, q-1\}$, then for any DFAO $M = (Q, \Sigma, \delta, q_0, \Delta, \eta)$ the sequence $\delta(q_0, \langle c(n) \rangle_q)$ is eventually periodic. This is true, since M has only finitely many states, thus there are $n_1 < n_2$ such that $\delta(q_0, \langle c(n_1) \rangle_q) = \delta(q_0, \langle c(n_2) \rangle_q)$. From n_2 on, the sequence $\delta(q_0, \langle c(n) \rangle_q)$ runs through the same cycle as before, since it follows the same arrows. When representing $q^n - 1$ in base q, this is the string consisting of n times the digit $q-1$, thus the above observation shows that $(a_{q^n-1})_{n \in \mathbb{N}}$ is eventually periodic if $(a_n)_{n \in \mathbb{N}}$ is q-automatic.

We use Christol's theorem to show that Carlitz π, i.e.,

$$\pi_q = \prod_{k=1}^{\infty} \left(1 - \frac{X^{q^k} - X}{X^{q^{k+1}} - X}\right),$$

(compare Chapter I.1) is transcendental. Our proof is due to Allouche [All93], see also [Sha04]. Suppose that π_q is algebraic. Since in this case π_q' would also be algebraic, the series

$$\frac{\pi_q'}{\pi_q} = \sum_{k=1}^{\infty} \left(\frac{1}{1 - \frac{X^{q^k} - X}{X^{q^{k+1}} - X}}\right) \left(\frac{X^{q^{k+1}} - X^{q^k}}{\left(X^{q^{k+1}} - X^{q^k}\right)^2}\right) = \left(\sum_{k=1}^{\infty} \frac{1}{X^{q^k} - X}\right) - \frac{1}{X^q - X}$$

would be algebraic. Thus, to prove that π_q is transcendental we have to show that the so called **bracket series**

$$\sum_{k=1}^{\infty} \frac{1}{X^{q^k} - X}$$

is transcendental. We write this as a power series in X^{-1}. We have

$$\begin{aligned}
\sum_{k=1}^{\infty} \frac{1}{X^{q^k} - X} &= X^{-1} \sum_{k=1}^{\infty} \frac{1}{X^{q^k - 1}\left(1 - \left(\frac{1}{X}\right)^{q^k - 1}\right)} = X^{-1} \sum_{k=1}^{\infty} \frac{1}{X^{q^k-1}} \sum_{n=0}^{\infty} \left(\frac{1}{X}\right)^{n(q^k-1)} \\
&= X^{-1} \sum_{n,k=1}^{\infty} \left(\frac{1}{X}\right)^{n(q^k-1)} = X^{-1} \sum_{m=1}^{\infty} \left(\frac{1}{X}\right)^m \sum_{\substack{k,n \geq 1 \\ n(q^k-1)=m}} 1 \\
&= X^{-1} \sum_{m=1}^{\infty} \left(\frac{1}{X}\right)^m c(m)
\end{aligned}$$

with

$$c(m) = \sum_{\substack{k \geq 1 \\ q^k - 1 \mid m}} 1.$$

We need to show that $c(m) \bmod p$ is not q-automatic and do this by showing that $c(q^n - 1) \bmod p$ is not eventually periodic. We have

$$c(q^n - 1) = \sum_{\substack{k \geq 1 \\ q^k - 1 | q^n - 1}} 1 = \sum_{\substack{k \geq 1 \\ k | n}} 1 = \tau(n).$$

Suppose that $(\tau(n))_{n \in \mathbb{N}}$ is eventually periodic modulo p and let n_0 denote the length of the pre-periodic part and L the length of the period. Choose k_0 such that $1 + k_0 L \geq n_0$ and $1 + k_0 L \in \mathbb{P}$. This is possible due to Dirichlet's theorem on primes in arithmetic progressions (Theorem 0.5.2). Then

$$\tau\left((1 + k_0 L) + k_0(1 + k_0 L)L\right) \equiv \tau(1 + k_0 L) \bmod p.$$

But $1 + k_0 L$ is a prime, say, ω and $1 + k_0 L + k_0(1 + k_0 L)L = \omega^2$, thus this would imply

$$3 = \tau(\omega^2) \equiv \tau(\omega) = 2 \bmod p,$$

and this contradiction shows that $c(m) \bmod p$ is not q-automatic, i.e., π_q is transcendental.

There are more related ways to consider transcendence with methods from automata theory, see [Tha, Chr79]. Chomsky and Schützenberger have shown yet another result relating formal languages and algebraic power series, see [CS63]. More connections between languages, automata, and Laurent series can be found in [Fir10]. For more about automatic sequences (partially with applications in number theory) see [Lot05, AS03].

As we already noted we will deal with another connection between computer science and number theory in Chapter II.1. There, we will consider the following problem: Given a language L, find the smallest subset $M \subset L$ such that any word $w \in L$ can be generated by some word $\widetilde{w} \in M$. Here a word w can be generated by $\widetilde{w}$ if we can obtain $\widetilde{w}$ from w by deleting some of its symbols (i.e., in our case, digits). In general it is hard to find such a set M for arbitrary languages L. However, for some languages defined via arithmetic conditions this is possible, and in some cases there is even an algorithm that yields M.

Other connections between number theory and computer science can be found in [Sha04] and the references mentioned there.

I.12

Algebraic Geometry

In view of what we have seen by now, this chapter can be seen as some kind of roundup, since here we will see a lot of concepts and problems we already discussed, cf. Figure I.C.1.

In Chapter I.7, we have seen that we can solve geometric problems (namely, the search for Pythagorean triples or congruent numbers) with algebraic or arithmetic methods. In this chapter, we take a look at the reverse:

Given a polynomial equation $f(x_1, \ldots, x_n) = 0$ (where $f \in \mathbb{Z}[x_1, \ldots, x_n]$), we are interested in rational or integer solutions. To find solutions, we associate to this equation a geometric object and use means of algebraic geometry to solve the equation or to find the number of solutions of this equation.

We do this by considering the algebraic curve defined by f. The maybe easiest nontrivial example is the Pythagorean equation $x^2 + y^2 = z^2$. As indicated in Paragraph 0.5.1, there are many ways to prove the characterization of Pythagorean triples. The method in which we are interested here is the geometric method: We normalize the equation and search for rational points on the curve $x^2 + y^2 = 1$. This can be done by finding one rational point (x_0, y_0) and consider lines with rational slopes through this point. The intersection of this line with the circle $x^2 + y^2 = 1$ then is another rational point (x_1, y_1), cf. Figure I.12.1.

In fact, this method can be generalized to arbitrary curves of degree 2, provided that the curve has at least one rational point, see [ST94]. It can be checked

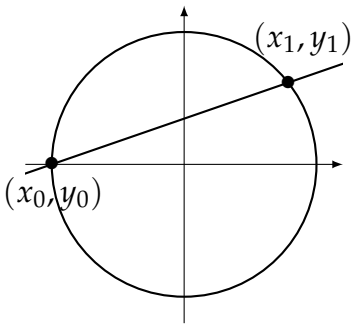

Figure I.12.1: Finding Pythagorean triples.

with the Hasse-Minkowski theorem (Theorem I.1.4) and Hensel's lemma (Theorem I.1.3) whether a curve has a rational point to start with.

For curves C of degree $m > 3$, Falting's theorem (see [Fal83]) states that C has only finitely many rational points. Therefore we take a look at curves of degree 3. Unless stated otherwise, all proofs can be found in [Sil08].

Definition I.12.1 An **elliptic curve** is a smooth projective algebraic curve of degree 3.

We can study elliptic curves for arbitrary fields K. We say that the elliptic curve E is **defined over K** if all coefficients of the minimal polynomial of E belong to K. If K does not have characteristic 2 or 3, any curve of degree 3 with a rational point is defined by an equation of the form $y^2 = x^3 + ax + b$ with $a, b \in K$ (a so-called **Weierstraß equation**) together with a point $\mathcal{O}$ at infinity. This equation defines a smooth (and thus elliptic) curve if the **discriminant** $\Delta_E = 4a^3 + 27b^2$ is nonzero (in the literature there are different definitions of Δ_E, but they only differ by a factor $\pm k^2$). If K has characteristic 2 or 3, there are other simple equations for elliptic curves, see [Sil08]. For our purpose, we will mostly have $K = \mathbb{Q}$. Figure I.12.2 shows some elliptic curves (these plots are not true to scale, they just serve as an indicator of how elliptic curves can look like).

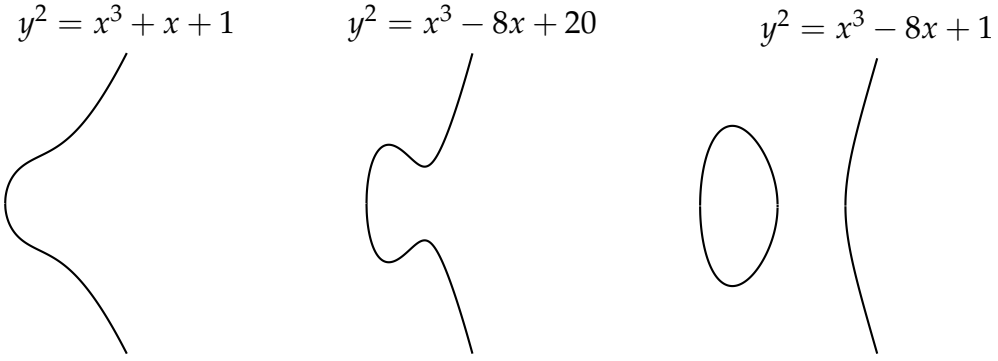

Figure I.12.2: Some elliptic curves.

Note that an elliptic curve does not look like an ellipse. The name has been given because these curves arise in the computation of the arc length of an ellipse, see [ST94, Exercise 1.16].

If E is an elliptic curve over $\mathbb{Q}$, we can define an addition law on E so that the rational points on E, denoted $E(\mathbb{Q})$, form an abelian group (more generally, this is true for any field K and we denote the corresponding group by $E(K)$). There is an algebraic expression of the addition law which is rather complicated, a better way to describe the addition is the geometric way, see Figure I.12.3.

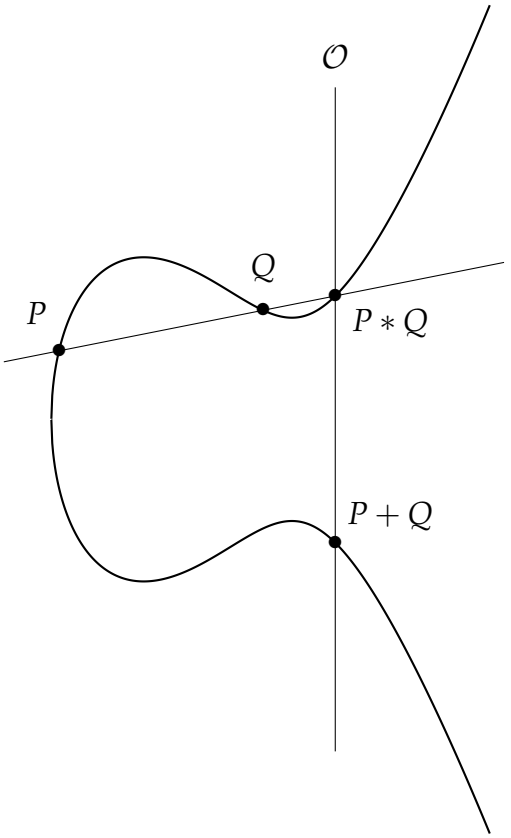

Figure I.12.3: The addition law for an elliptic curve.

Elliptic curves have been the objects of intense recent studies. Before looking at applications to problems we have described, we just mention two results that are important for our interest (i.e., looking for solutions of equations), the **Mordell-Weil theorem** and **Siegel's theorem**:

Theorem I.12.2 (Mordell-Weil theorem) *Let E be an elliptic curve defined over $\mathbb{Q}$. Then the group $E(\mathbb{Q})$ is finitely generated, i.e., $E(\mathbb{Q})$ has finite rank.*

Theorem I.12.3 (Siegel's theorem) *Let E be an elliptic curve defined over $\mathbb{Q}$. Then there are only finitely many integer points on E.*

Unfortunately, the Mordell-Weil theorem does not tell us how to find the rank of E. But there are results about rational points of finite order (the Nagell-Lutz theorem), results on bounds of the coefficients of rational points, as well as results about the torsion group (Mazur's theorem) that can help to determine the rank, cf. [LR11]. For more basic information about elliptic curves see [ST94, Sil08], for advanced topics see [Sil99].

With elliptic curves we can handle a lot of problems concerning the search for integer solutions of equations of degree 3. We emphasize three of the problems which we have seen already in earlier chapters: The Mordell equation, congruent numbers and Fermat's last theorem.

The Mordell equation $y^2 = x^3 + k$ with $k \in \mathbb{Z}$ defines an elliptic curve if $k \neq 0$. In [GPZ98], the authors use this fact to determine the solutions of Mordell's equation for $|k| \leq 10\,000$, see also [NTW]. Unfortunately, it seems that the number of solutions has to be computed for each k individually, there is no formula that gives the integer points on the elliptic curve $y^2 = x^3 + k$.

The problem of congruent numbers can also be investigated with the help of elliptic curves, see [Conh, Kob93]. For a natural number n, consider the elliptic curve $y^2 = x^3 - n^2x$. We have the following correspondence:

Theorem I.12.4 *Let $n \in \mathbb{N}$. Then the maps*

$$(a,b,c) \mapsto \left(\frac{nb}{c-a}, \frac{2n^2}{c-a}\right), \qquad (x,y) \mapsto \left(\frac{x^2-n^2}{y}, \frac{2nx}{y}, \frac{x^2+n^2}{y}\right)$$

define a bijection between the sets

$$\left\{(a,b,c) \in \mathbb{Q}^3 : a^2 + b^2 = c^2, \frac{1}{2}ab = n\right\} \text{ and } \left\{(x,y) \in \mathbb{Q}^2 : y^2 = x^3 - n^2x, y \neq 0\right\}.$$

This tells us that rational side lengths of rectangular triangles with area n correspond to nontrivial rational points on the elliptic curve $y^2 = x^3 - n^2x$. Examining this, Tunnel found necessary conditions for n to be congruent, see [Conh]. The question whether these conditions are also sufficient relates to an important open conjecture, the **Birch and Swinnerton-Dyer conjecture**. To state this conjecture, we associate an L-function to the elliptic curve E, the **Hasse-Weil *L*-function** $L(E,s)$ (for an exact definition and properties see [Sil99]). The weak form of the Birch and Swinnerton-Dyer conjecture is:

Conjecture I.12.5 (Birch and Swinnerton-Dyer conjecture) *Let E be an elliptic curve defined over $\mathbb{Q}$. Then $L(E,1) = 0$ if and only if E has infinitely many rational points.*

There is also a strong form of this conjecture that relates the rank of $E(\mathbb{Q})$ with the Taylor expansion of $L(E,1)$. The Birch and Swinnerton-Dyer conjecture is one of the Millennium Problems, cf. [CJW06]. For an introduction to the Birch and Swinnerton-Dyer conjecture see [Zag91, Wil06, Hus04].

If the Birch and Swinnerton-Dyer conjecture is true, then the conditions of Tunnel are also sufficient for n to be a congruent number, see [Kob93]. Further the Birch and Swinnerton-Dyer conjecture implies that any natural number n with $n \equiv 5, 6, 7 \bmod 8$ is congruent, see [LR11]. Further equivalences to the congruent number problem can be found in [Kob93].

Under certain conditions there is a connection between the Hasse-Weil L-function $L(E,s)$ for an elliptic curve E defined over some field K and the Hecke L-function $L(s,\chi)$ attached to a Hecke character of $J_{K'}$ for some field extension K' of K. This has been proved by Deuring [Deu53], see also [Sil99]. This connection can be used to transfer the functional equation from $L(s,\chi)$ to $L(E,s)$.

Finally, we mention what is possibly most prominent use of elliptic curves for Diophantine equations, namely Fermat's last theorem:

Theorem I.12.6 (Fermat's last theorem) *For $n \geq 3$, the equation $x^n + y^n = z^n$ has no nontrivial integer solutions.*

The steps in the proof of Fermat's last theorem are rather complicated and not at all straight-forward, therefore we will just mention the results one needs in order to prove Fermat's last theorem. For some more information see [Fal95, Hus04].

First we need the notion of modular curves. There are many equivalent definitions of modular curves. We will state two of them.

If E is an elliptic curve over $\mathbb{Q}$ whose Weierstraß equation has integer coefficients and p is a prime, we can define a curve $\widetilde{E}$ over $\mathbb{F}_p$ by reducing all coefficients of the Weierstraß equation of E modulo p. If $\widetilde{E}$ is not singular, it is again an elliptic curve and we set $a_p = p + 1 - \left|\widetilde{E}(\mathbb{F}_p)\right|$. An elliptic curve is called **modular** if there is an eigenform $f \in \mathscr{S}_2(N)$ (for some $N \in \mathbb{N}$) such that for any $p \nmid N$ we have $T_p(f) = a_p f$ (compare Chapter I.10). In this case the integers a_p are exactly the Fourier coefficients of f.

Equivalently, an elliptic curve E is called modular if there is an $f \in \mathscr{S}_2(N)$ (for some $N \in \mathbb{N}$) such that $L(E,s) = L(s,f)$. For other equivalent definitions of modular curves see [DS05, Sil08, Rib90].

Now the outline of the proof of Fermat's last theorem is as follows: Let p be an odd prime. If $(a, b, c) \in \mathbb{Z}^3$ is a nontrivial solution of Fermat's equation $x^p + y^p = z^p$, we associate to this solution the **Frey curve** $y^2 = x(x - a^p)(x + b^p)$. Ribet [Rib90] proved that this Frey curve is not modular. On the other hand, the **Taniyama-Shimura-Weil conjecture** states that every elliptic curve defined over $\mathbb{Q}$ is modular. As a consequence of this, the Taniyama-Shimura-Weil conjecture implies Fermat's last theorem.

Later, Wiles [Wil95] proved (with some flaws that were corrected in [TW95]) the Taniyama-Shimura-Weil conjecture, now known as the **modularity theorem**, in a special case.

Theorem I.12.7 (modularity theorem) *Every elliptic curve over $\mathbb{Q}$ is modular.*

Together with the result of Ribet, this proves Fermat's last theorem.

I.C

Conclusion

We have seen several topics and methods connecting number theory with various other mathematical areas. Of course, our treatment is to no extent complete: One could identify other areas and consider connections to those, as well as find other connections to the areas that we discussed. In this sense, the first part of this thesis can be seen as an introduction to the possibilities and the variety of number theory.

This presentation would be a nice basis for future work. One could pick an area and show other connections (I think it is impossible to show all connections to a given area, since the field of mathematics is too broad and its development is too fast). For some mathematical areas it is possibly hard to find connections to number theory (for example for category theory, as already mentioned in the introduction), but in most mathematical areas there are nice ones. One could also pick a certain topic (for example ζ-functions, modular forms, ...) and investigate the areas in which this topic is present.

Some more connections and topics that would be suitable for connections can be found in the literature mentioned in the introduction.

It is noteworthy that there are connections between most of the introduced concepts, i.e., these concepts are not only connections between number theory and one other area, but rather between number theory and more other areas. This is summarized in Figure I.C.1. Ironically, the only reason why this graph is

not connected is graph theory. Of course this can easily be repaired: In the next part of this thesis we will see a connection between graph theory and number theory that involves sumsets.

Maybe the graph in Figure I.C.1 cannot be seen as a "map of math" (like Figures 0.0.1 and 0.0.2), since too many areas and connections between other areas are missing. But we can view this as (part of) a "map of the neighbourhood of number theory".

In the following part of this thesis we will see some more connections between number theory and some mathematical areas.

All of the following connections have already been mentioned in the respective chapters: In Chapter II.1 we deal with the concept of minimal sets mentioned in Chapter I.11. In Chapter II.2 we use number theoretic methods to obtain properties of graphs mentioned in Chapter I.5. In Chapter II.3 we use algebraic methods to solve equations that are useful in number theory, as briefly described in Chapter I.8. In Chapter II.4 we use number theoretic and analytic tools to solve a puzzle that can be modeled with linear algebra, cf. Chapters I.1 and I.6.

All results in the subsequent part are new results and (unless indicated otherwise) my own research results. These results have been published in the articles mentioned in the respective chapters.

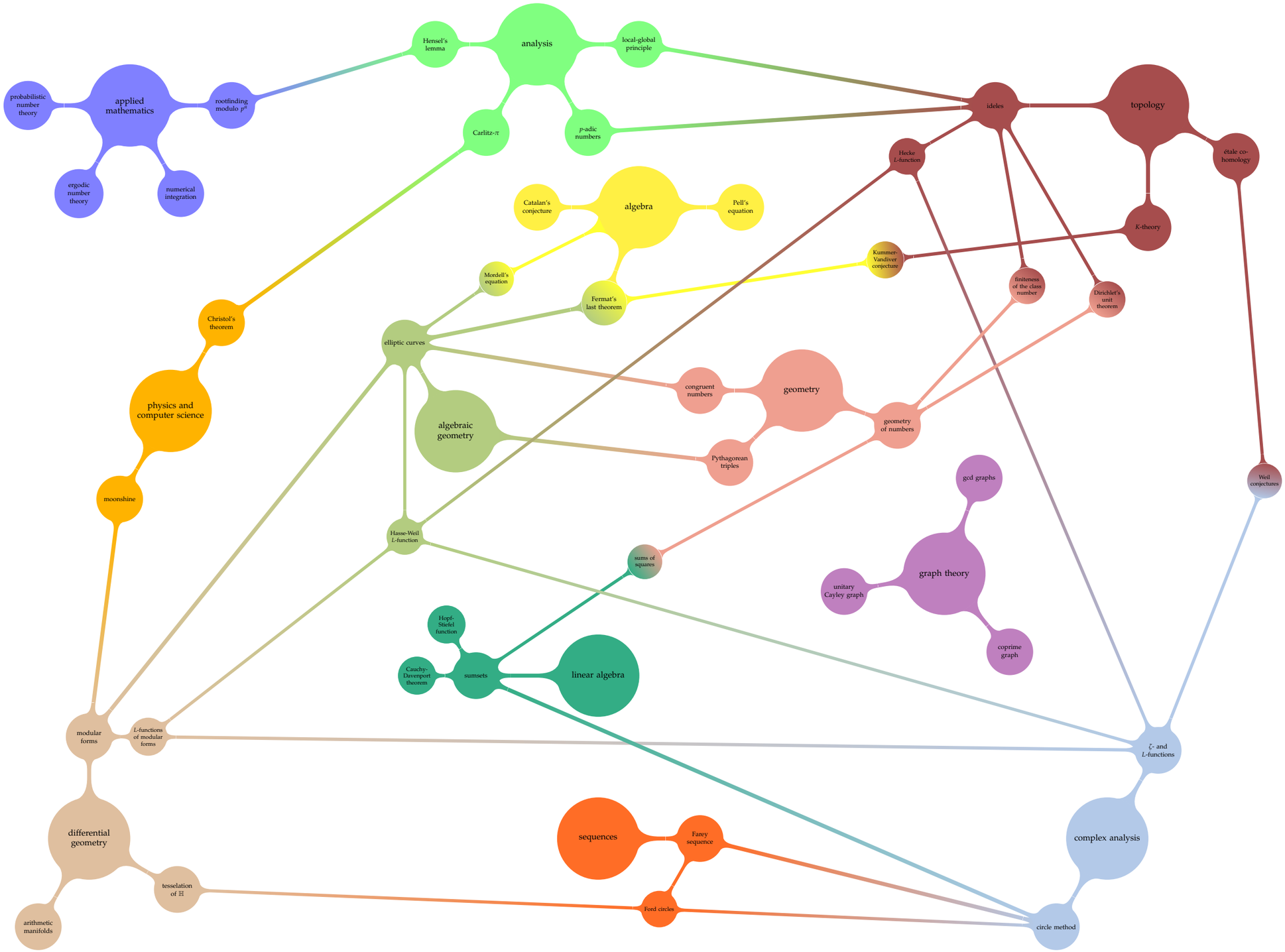

Figure I.C.1: Connections between the areas and topics discussed in the first part of this thesis.

Part II - Original Work

II.1

Minimal Sets

In this chapter we examine a connection between number theory and theoretical computer science. More precisely, we consider a concept in the theory of formal languages, namely the concept of minimal sets and comparable elements. We will apply this to languages defined over the set of digits $\{0, \dots, 9\}$, thus computing minimal sets of natural numbers written in decimal expansion.

The results have been published in [Kre15b] and [BKS17], where the second article is joint work with Jörn Steuding and Ioulia Baoulina.

II.1.1 Introduction

In 2001, Shallit [Sha00a] introduced a problem concerning decimal digits of some sets of natural numbers. To state the problem, let us first fix some notation (partially from [Sha00a]).

We consider the natural numbers with respect to their unique decimal expansion, where each $n \in \mathbb{N}$ is given by $n = \sum_{j=0}^{k} \alpha_j 10^j$ for some natural number k and digit sequence $\langle n \rangle_{10} = \alpha_k \alpha_{k-1} \dots \alpha_0$ with $\alpha_k \neq 0$. Since we will almost only deal with the base 10 representation (the only exception is in Section II.1.7), we will, with slight abuse of notation, use n instead of $\langle n \rangle_{10}$. For other bases all concepts can be defined analogously.

If we refer to the decimal representation, we will sometimes call $a_k a_{k-1} \cdots a_1$ a **string**. Let x and y be two strings. We say that x is a **subsequence** of y if $x = y$ or if we can obtain x by deleting some digits of y. Equivalently we will say that y can be **generated** from x. If x is a subsequence of y, we will write $x \triangleleft y$. Otherwise, we write $x \not\triangleleft y$. For example, we have $134 \triangleleft 918234 \triangleleft 98188293894$ but $123 \not\triangleleft 43021$. If for two strings x, y either $x \triangleleft y$ or $y \triangleleft x$, we call x and y **comparable**, and otherwise **incomparable**. If M is a set such that any two strings in M are incomparable, we call M **not truncatable**.

If x and y are two strings, let $x * y$ denote the **concatenation** of x and y, i.e., the string that has first the digits from x and then the digits from y. Sometimes we will omit the asterisk. For two sets $M, L \subset \mathbb{N}$ we define

$$M * L := \{z \in \mathbb{N} : z = x * y, x \in M, y \in L\}.$$

Again, we will omit the asterisk sometimes and write ML instead of $M * L$. If M contains only one element, i.e., $M = \{m\}$ for some m, we even omit the braces and write mL instead of $\{m\}L$.

For $x, k \in \mathbb{N}$ define inductively $x^{*k} := x^{*(k-1)} * x, x^{*1} := x$ and let $\{x\}^*$ denote the set

$$\{x\}^* = \{x^{*k} : k \in \mathbb{N}_0\}.$$

Here x^{*0} means that x does not occur (this can be made precise with the so-called **empty word**, but we will not need this). For example,

$$2\{13\}^*4\{5\}^* = \{24, 2134, 245, 21345, 213134, 2455, 2131345, 213455, 21313455, \ldots\}.$$

If $M \subset \{0, 1, \ldots, 9\}$, we set

$$M^* := \{x \in \mathbb{N}_0 : (d \triangleleft x, d \in \{0, \ldots, 9\}) \Rightarrow d \in M\},$$

i.e., M^* contains only the natural numbers all of whose digits are in M. For $n \in \mathbb{N}$, we let $\#n$ denote the number of digits of n. Let $M \subset \mathbb{N}_0$ and $I \subset \mathbb{N}$. Then we define the set M^{*I} by

$$M^{*I} := \{x \in M^* : \#x \in I\}.$$

For example we have $123\{4\}^{*\{5\}} = 12344444$ and

$$\{1,5\}^{*\{2n:n\in\mathbb{N}\}} = \{11, 15, 51, 55, 1111, 1115, 1151, 1511, 5111, 1155, 5511, 1551, 5115, 1515, 5151, 1555, 5155, 5515, 5551, 5555, 111111, \ldots\}.$$

If $L \subset \mathbb{N}$, we will call

$$\langle L \rangle := \{x \in \mathbb{N} : \exists y \in L \text{ such that } y \lhd x\}$$

the **generated set** of L. It is clear that $\langle \{1,2,3,4,5,6,7,8,9\} \rangle = \mathbb{N}$.

The main problem for this chapter is as follows: Given a set $M \subset \mathbb{N}$, find the smallest possible set $A \subset M$ such that for all $m \in M$ there exists $a \in A$ with $a \lhd m$. As a matter of fact, the set

$$\mathcal{S}(M) := \{m \in M : \{n \in M : n < m, n \lhd m\} = \emptyset\}$$

solves this problem: Clearly, every element of M contains some element of $\mathcal{S}(M)$ as a subword. The set $\mathcal{S}(M)$ is said to be the **minimal set** of M and its elements are called **minimal**.

Note that a subset of M consisting of pairwise incomparable elements is in general not contained in $\mathcal{S}(M)$ (as the example $M = \mathbb{N}$ and the subset of all two digit numbers shows). However if for any element x in a set $A \subset M$ of pairwise incomparable elements there is no $y \in M, y \neq x$ with $y \lhd x$, then A is indeed contained in $\mathcal{S}(M)$. We shall use this idea quite often in the sequel.

Remarkably, $\mathcal{S}(M)$ is finite for every $M \subset \mathbb{N}$. This follows from a celebrated theorem from the theory of formal languages, the so-called lemma of Higman [Hig52] (see also [SS83]).

In [Sha00a], Shallit computed the minimal set for the primes and the composite natural numbers, and made a conjecture about the powers of 2. He showed that

$$\mathcal{S}(\mathbb{P}) = \{2,3,5,7,11,19,41,61,89,409,449,881,991,6469,9001,9049,9649,\\ 9949,60649,666649,946669,60000049,66000049,66600059\}$$

and

$$\mathcal{S}(C) = \{4,6,8,9,10,12,15,20,21,22,25,27,30,32,33,35,50,51,52,55,57,70,\\ 72,75,77,111,117,171,371,711,713,731\},$$

where C denotes the composite numbers, i.e., $C = \mathbb{N}\backslash(\mathbb{P} \cup \{1\})$. Moreover, he conjectured that

$$\mathcal{S}(\{2^n : n \in \mathbb{N}_0\}) = \{1,2,4,8,65536\}$$

and he observed that this conjecture is true if every number 16^m with $m \geq 4$ has at least one digit from $\{1, 2, 4, 8\}$. This type of problem seems to be difficult to tackle. Since $\log_{10}(2)$ is irrational, Theorem I.3.6 (and the comment afterwards) yields that the sequence $\{n \log_{10}(2)\}$ of the fractional parts of $n \log_{10}(2)$ is uniformly distributed. Hence the powers of 2 are distributed according to **Benford's law** (which implies that the probability that a power of 2 has leading digit k is exactly $\log_{10}(1 + \frac{1}{k})$, cf. [Ste] for details). Nevertheless, this just indicates - in an appropriate probabilistic framework - that a power of 2 without any digit from $\{1, 2, 4, 8\}$ must be a very rare event.

This conjecture already shows that it may be hard to determine the minimal set for a given infinite set.

In a more recent work, Bright, Devillers, and Shallit [BDS16] computed the minimal set for the primes in base b representation for $2 \leq b \leq 30$ (for some b only partially), the case $b = 10$ being the above fact already handled in [Sha00a].

Gruber, Holzer, and Kutrib studied the problem from the viewpoint of (theoretical) computer science, concerning (effective) constructions and decidability of problems that involve so-called Higman-Haines sets (see [GHK07, GHK09]).

Although we know that the minimal set is always finite, known proofs of this fact are not constructive. As a consequence, it seems to be difficult to compute $\mathcal{S}(M)$ for a given set M in general. An idea for an algorithm would be to strike out all natural numbers in M that can be generated by a smaller element of M and therefore do not belong to the minimal set. But since M is infinite, we cannot (in finite time) check all elements from M without using special properties of the set M. Actually, Gruber, Holzer, and Kutrib showed that the problem of determining the minimal set for a given set M is in general unsolvable, see [GHK07].

In this chapter we will examine some more examples of minimal sets and some conceptual approaches.

In Sections II.1.2 and II.1.3 we will first present some new results about sets defined via arithmetic conditions, for example the set of natural numbers that can be written as the sum of two squares and sets of natural numbers that are values of arithmetic functions. Additionally we will get a result about repdigits that are sums of squares.

In Section II.1.4 we consider a special class of sets, namely congruence classes. We will compute the minimal sets for a few first examples and develop an algorithm that gives the minimal set for "congruence class like" sets.

In these cases the minimal sets can be determined easily. Whereas in the first two sections the reason for this is that the examined sets contain many digits, in Section II.1.4 we exploit the fact that the elements of the sets are in some sense well-distributed.

In the next two sections we will examine conceptual results. In Section II.1.5 we consider basic set operations, and in Section II.1.6 we will deal with heuristics and measures. In these sections we will see that conceptual approaches are unlikely to be successful. We will end our treatment with some remarks about perfect numbers in Section II.1.7.

II.1.2 Sums of squares and repdigits

In this section we look at some subsets of $\mathbb{N}$ for which the minimal sets can be determined easily.

In all explicitly given lists here and later, one first has to check that all the elements belong to the set M one considers, and that there are no $x, y \in \mathcal{S}(M)$ with $x \neq y, x \triangleleft y$ (i.e., that $\mathcal{S}(M)$ is not truncatable). This is always easy (but can get exhausting when $\mathcal{S}(M)$ has a lot of elements) and is mostly left to the reader.

Theorem II.1.1 *Let*

$$\boxed{3} := \{n \in \mathbb{N} : n = x^2 + y^2 + z^2 \text{ for some } x, y, z \in \mathbb{N}_0\}.$$

Then

$$\mathcal{S}(\boxed{3}) = \{1, 2, 3, 4, 5, 6, 8, 9, 70, 77\}.$$

Proof. Clearly $1, 2, 3, 4, 5, 6, 8, 9$ belong to $\mathcal{S}(\boxed{3})$. Since $7 \notin \boxed{3}$, 70 and 77 also belong to $\mathcal{S}(\boxed{3})$. If $n \in \boxed{3}$ is different from $1, 2, 3, 4, 5, 6, 8, 9, 70, 77$, then either $d \triangleleft n$ for $d \in \{1, 2, 3, 4, 5, 6, 8, 9\}$ hence n cannot belong to $\mathcal{S}(\boxed{3})$, or n consists only of the digits 0 and 7. In the latter case either $70 \triangleleft n$ or $77 \triangleleft n$, and thus $n \notin \mathcal{S}(\boxed{3})$. q.e.d.

The result for numbers that can be written as the sum of two squares is harder. We will need (a part of) the following lemma:

Lemma II.1.2

1. *Let $x \in \{7\}^*$. Then x cannot be written as the sum of two squares.*

2. *Let $x \in \{3\}^*$. Then x is the sum of two squares if and only if $x = 333$.*

Proof.

1. (partially from [Tro]) Let σ'_k denote the number that contains exactly k times the digit 7. Then

$$\sigma'_k = 7 \cdot \sum_{i=0}^{k-1} 10^i = 7 \cdot \frac{10^k - 1}{10 - 1} = \frac{7}{9}(10^k - 1).$$

Since 9 is a square, σ'_k is the sum of two squares if and only if $\sigma_k := 9\sigma'_k$ is the sum of two squares. We start with two claims.

- Claim 1: If σ'_k is the sum of two squares, then k is even. Since $7|\sigma'_k$ and $7 \equiv 3 \bmod 4$, 7 has to divide the number $\underbrace{11\ldots11}_{k}$. If testing for divisibility by 7, we can split the number into blocks of 3 digits and alternately add and subtract them. Any two full blocks will cancel while non-full blocks will not give a multiple of 7 (one can easily check this). It follows that $7^2|\sigma'_k$ if and only if $6|k$. In particular, k has to be even.
- Claim 2: If $k > 1$ and σ_{2k} can be written as the sum of 2 squares, then σ_k can be written as the sum of two squares. We have

$$\sigma_{2k} = 7(10^{2k} - 1) = 7(10^k - 1)(10^k + 1) = \sigma_k(10^k + 1).$$

Since $\gcd(10^k - 1, 10^k + 1) = 1$, we get $\gcd(\sigma_k, 10^k + 1) \in \{1, 7\}$. Since σ_{2k} can be written as the sum of two squares, we have $2|v_p(\sigma_{2k})$ for every $p \equiv 3 \bmod 4$. For $p \neq 7$ we have $p \nmid \gcd(\sigma_k, 10^k + 1)$, so $2|v_p(\sigma_k)$. For $p = 7$ either $2|v_7(\sigma_k)$ (if $\gcd(\sigma_k, 10^k + 1) = 1$) or $2 \nmid v_7(\sigma_k)$ (if $\gcd(\sigma_k, 10^k + 1) = 7$). So either σ_k or $7\sigma_k$ can be written as the sum of two squares. But for $k > 1$ we have $7\sigma_k \equiv 3 \bmod 4$, thus σ_k can be written as the sum of two squares.

Now consider σ_k with $k = 2^m l$ such that $2 \nmid l, m \in \mathbb{N}_0$. If $m = 0$, σ_k is not the sum of two squares, since k is odd. For $l \neq 1$, the second claim yields that σ_l is the sum of two squares if σ_k is the sum of two squares. Since this is false, σ_k is not the sum of two squares. Now let $l = 1$. Then, again due to the second claim, if σ_k is the sum of two squares, so is $\sigma_2 = 9 \cdot 77$. Since 77 is not the sum of two squares, σ_k is not the sum of two squares. This completes the proof.

2. Let δ'_k denote the natural number that contains exactly k times the digit 3, i.e., $\delta'_k = \frac{3}{9}(10^k - 1)$. Since $3|\delta'_k$ and $3 \equiv 3 \bmod 4$, 3 has to divide the number $\underbrace{11\ldots11}_{k}$. This is the case if and only if $3|k$. We want to show that $\delta_{3k} := 9\delta'_{3k}$ cannot be the sum of two squares for $k > 1$. Suppose that δ_{3k} is the sum of two squares. We have

$$\delta_{3k} = 3(10^{3k} - 1) = 3(10^k - 1)(10^{2k} + 10^k + 1) = \delta_k(10^{2k} + 10^k + 1).$$

Since $10^{2k} + 10^k + 1 - (10^k + 2)(10^k - 1) = 3$, the greatest common divisor of $10^k - 1$ and $10^{2k} + 10^k + 1$ is either 1 or 3. But $10^{2k} + 10^k + 1 \equiv 0 \bmod 3$ and $10^k - 1 \equiv 0 \bmod 3$, thus $\gcd(10^k - 1, 10^{2k} + 10^k + 1) = 3$. More precisely, we have $10^{2k} + 10^k + 1 \equiv 3 \bmod 9$, therefore $\gcd(\delta_k, 10^{2k} + k + 1) = 3$. Since δ_{3k} is the sum of two squares, we have $2|v_p(\delta_{3k})$ for all p with $p \equiv 3 \bmod 4$. For $p \neq 3$ we have $p \nmid \gcd(\delta_k, 10^{2k} + 10^k + 1)$, so $2|v_p(\delta_k)$. For $p = 3$ we get $2 \nmid v_3(\delta_k)$. It follows that $3\delta_k$ is the sum of two squares. But for $k > 1$ we have $3\delta_k \equiv 3 \bmod 4$, so this is a contradiction.

q.e.d.

The result about the minimal set for the set of numbers that can be written as a sum of two squares will also rely on the following conjecture:

Conjecture II.1.3 *Let*

$$A := \left\{ n \in \mathbb{N} : n \in 700\{7\}^{*\{66k+61, k\in\mathbb{N}\}} \text{ and } n = x^2 + y^2 \text{ for some } x, y \in \mathbb{N}_0 \right\}.$$

Then $A = \emptyset$.

Theorem II.1.4 *Let*

$$\boxed{2} := \{n \in \mathbb{N} : n = x^2 + y^2 \text{ for some } x, y \in \mathbb{N}_0\}.$$

If Conjecture II.1.3 is true, then

$$\mathcal{S}(\boxed{2}) = \{1, 2, 4, 5, 8, 9, 36, 37, 73, 333, 666, 676, 677, 706, 776, 60633, 77077, 7000777\}.$$

If Conjecture II.1.3 is not true and A is defined as in Conjecture II.1.3, then

$$\mathcal{S}(\boxed{2}) = \{1, 2, 4, 5, 8, 9, 36, 37, 73, 333, 666, 676, 677, 706, 776, 60633, 77077, 7000777, \min_{a\in A} a\}.$$

Proof. We will use the fact that the set $\boxed{2}$ is closed under multiplication. This follows from the identity

$$(a^2+b^2)(c^2+d^2) = (ac+bd)^2 + (ad-bc)^2.$$

Let $x \in \boxed{2}$ with length at least 2. If x contains the digit $1, 2, 4, 5, 8$, or 9, then $x \notin \mathcal{S}(\boxed{2})$. So let $x \in \{3,6,7\}\{0,3,6,7\}^*$. If the last digit of x is 0, then $x \in \boxed{2}$ is equivalent to $\frac{x}{10} \in \boxed{2}$ due to the two-squares theorem (Theorem 0.5.10). Since $\frac{x}{10} \lhd x$, x cannot be in $\mathcal{S}(\boxed{2})$. Hence the last digit of x cannot be 0. We consider the last two digits of x.

- If the last two digits are $37, 73$, or 36, we are done, since these are in $\mathcal{S}(\boxed{2})$.
- If the last two digits are $63, 67, 03$, or 07, x is congruent 3 modulo 4 and thus not in $\boxed{2}$.

There remain the cases $33, 76, 06, 66$, and 77.

33: Since $33 \notin \boxed{2}$, x has length at least three. If $7 \lhd x$, we are done, since $73 \lhd x$. If there is another 3 in x, we are also done, since $333 \lhd x$. So suppose that $x \in 6\{0,6\}^*33$. Since x is divisible by 3, it also has to be divisible by 9 due to the two-squares theorem. So 6 has to occur m times, with $m \equiv 2 \bmod 3$. If $m > 2$, we are done since then $666 \lhd x$, so suppose $m = 2$. Since 6633 is not the sum of two squares, there has to be a 0 in x. We have $x \in 6\{0\}^*6\{0\}^*33$. If there is a 0 in the first place, we have $60633 \lhd x$, so let $x \in 66\{0\}^*33$. But then $11|x$ and the alternating digit sum of $\frac{x}{11} \in 6\{0\}^*3$ ist either 3 or 9, which are both not divisible by 11. So we have $11^2 \nmid x$, thus $x \notin \boxed{2}$.

76: Since $76 \notin \boxed{2}$, x has at least length three. If $3 \lhd x$, we have $37 \lhd x$. If not, x contains another 6 or 7, and then $676 \lhd x$ or $776 \lhd x$.

06: If x contains a 3, we are done. If x contains a 7, we have $706 \lhd x$. So assume $x \in 6\{0,6\}^*6$. But then 6 has to occur with multiplicity $3k$, since otherwise $3|x$ and $3^2 \nmid x$. Thus $666 \lhd x$.

66: We have $66 \notin \boxed{2}$, and if x has at least three digits, then $666 \lhd x$ or $36 \lhd x$ or $x \in 7\{0,7\}^*66$. In the latter case, if x contains no 0, then x cannot be written as the sum of two squares, since otherwise $\frac{1}{2}x \in 3\{8\}^*3$ could be written as the sum of two squares. But $\frac{1}{2}x \equiv 3 \bmod 4$, thus x has to contain a 0 and we get $706 \lhd x$.

77: If x contains a 3 or 6, we are done, so let $x \in 7\{0,7\}^*77$. If x does not contain a 0, then $x \notin \boxed{2}$ due to Lemma II.1.2. If $77077 \triangleleft x$, we are done, so let $x \in 7\{0\}^*\{7\}^*77$. Let k denote the number of zeros and m denote the number of sevens in $\{0\}^*\{7\}^*$. We know that $k \geq 1$. Since $7|x$ we need to have $7^2|x$. Note that

$$\frac{x}{7} = 10^{m+k+2} + \frac{1}{9}(10^{m+2} - 1) \equiv 2 \cdot 3^m \cdot 3^k + 4 \cdot 3^{m+2} - 4 \bmod 7.$$

Thus $\frac{x}{7}$ is divisible by 7 if and only if

$$3^{m+1}(3^{k-1} - 1) \equiv 2 \bmod 7.$$

Modulo 7, there are 6 possibilities to write 2 as a product $2 = uv$, these are $(u,v) = (\pm 1, \pm 2), (\pm 2, \pm 1), (\pm 3, \pm 3)$. Each of these cases gives a congruence condition modulo 6 for k and m:

$$\begin{aligned}
(1,2) &\Rightarrow m \equiv 5 \bmod 6, \quad k \equiv 2 \bmod 6,\\
(2,1) &\Rightarrow m \equiv 1 \bmod 6, \quad k \equiv 3 \bmod 6,\\
(-1,-2) &\Rightarrow m \equiv 2 \bmod 6, \quad k \equiv 4 \bmod 6,\\
&\qquad (-2,-1) \text{ is not possible},\\
(3,3) &\Rightarrow m \equiv 0 \bmod 6, \quad k \equiv 5 \bmod 6,\\
(-3,-3) &\Rightarrow m \equiv 3 \bmod 6, \quad k \equiv 0 \bmod 6.
\end{aligned}$$

In the second case, the smallest possible x is 7000777 and this is the sum of two squares. So we only need to consider the remaining cases if $k < 3$ or $m < 1$, i.e., the first and the fifth case (note that $k > 0$). Further, in these cases we only need to consider the numbers x with $k = 2$ (in the first case) or $m = 0$ (in the fifth case).

If $m = 0$, we have $x \in 7\{0\}^*77$. Since $3|x$ but $3^2 \nmid x$ we have $x \notin \boxed{2}$. Let $x \in 700\{7\}^*77$. Since $m \equiv 5 \bmod 6$ is odd, we know that $11|x$ and if we write $m = 6k + 5$, we get

$$\frac{x}{11} = 637 \cdot 10^{6k+6} + \sum_{i=0}^{3k+2} 7 \cdot 10^{2i}.$$

The alternating digit sum of $\frac{x}{11}$ is $21k + 31$. Hence x is divisible by 11^2 if and only if $k \equiv 9 \bmod 11$, i.e., if $m \equiv 59 \bmod 66$. Thus the theorem follows.

q.e.d.

Recall that a **repdigit** is a natural number n all of whose digits (in decimal representation) are the same. If this digit is d, we call n a ***d*-repdigit**. A 1-repdigit is called **repunit**. Let us call a repdigit **true** if it has at least two digits. Then the previous results give us also the following theorem:

Theorem II.1.5

1. *All true d-repdigits with $d \in \{2, 3, 6, 7, 8\}$ are the sum of three squares.*
2. *The only true* 5*-repdigit that is not the sum of three squares is* 55*. The only true d-repdigits with $d \in \{1, 4, 9\}$ that are sums of three squares are* $11, 44, 99$*.*
3. *The only true repdigits that are the sum of two squares are* 333 *and* 666*.*
4. *No true repdigit is a square.*

II.1.3 Values of Arithmetic Functions

Now we consider sets of the type $M = \{f(m) : m \in \mathbb{N}\}$ where $f : \mathbb{N} \to \mathbb{N}$ is an arithmetic function.

In the proof of the subsequent result we frequently need to show that some natural numbers m do not occur as values of the Euler totient function. This can be done (at least for "small" m) easily by distinguishing cases on the prime factorization of m and those of a potential n with $\varphi(n) = m$, together with using the fact that $\varphi(n)$ is even for $n \geq 3$. Thus we will omit these arguments.

Theorem II.1.6 (due to I. Baoulina) *Let* $\varphi(\mathbb{N}) := \{\varphi(m) : m \in \mathbb{N}\}$*. Then*

$$\mathcal{S}(\varphi(\mathbb{N})) = \{1, 2, 4, 6, 8, 30, 70, 500, 900, 990, 5590, 9550, 555555555550\}.$$

Proof. Observe that the numbers above are pairwise incomparable. Further, as

$$\begin{array}{cccc} \varphi(1) = 1, & \varphi(3) = 2, & \varphi(5) = 4, & \varphi(7) = 6, \\ \varphi(16) = 8, & \varphi(31) = 30, & \varphi(71) = 70, & \varphi(625) = 500, \\ \varphi(1057) = 900, & \varphi(991) = 990, & \varphi(5591) = 5590, & \varphi(9551) = 9550, \end{array}$$

and

$$\varphi(555555555551) = 555555555550,$$

we see that

$$\{1, 2, 4, 6, 8, 30, 70, 500, 900, 990, 5590, 9550, 555555555550\} \subset \varphi(\mathbb{N}).$$

Note also that $3,5,7,9 \notin \varphi(\mathbb{N})$ since $\varphi(n)$ is even for $n \geq 3$. Now assume that $n \in \varphi(\mathbb{N})$ has at least two digits. If there exists $d \in \{1,2,4,6,8\}$ such that $d \lhd n$, then $n \notin \mathcal{S}(\varphi(\mathbb{N}))$. Suppose that n contains only the digits $0,3,5,7,9$. Since $\varphi(m)$ is even for $m > 2$, the last digit is 0. It is easy to see that $50,90 \notin \varphi(\mathbb{N})$. Now assume that n has at least three digits. If n contains the digits 3 or 7, then $30 \lhd n$ or $70 \lhd n$, respectively. Consequently, $n \notin \mathcal{S}(\varphi(\mathbb{N}))$. Next assume that n contains only the digits $0,5,9$. It is easily verified that $550,590,950 \notin \varphi(\mathbb{N})$. Suppose that n has at least four digits. Then we have the following possibilities:

- n contains at least two zeros. Then either $500 \lhd n$ or $900 \lhd n$.
- $99 \lhd n$. Then $990 \lhd n$.
- $559 \lhd n$. Then $5590 \lhd n$.
- $955 \lhd n$. Then $9550 \lhd n$.
- $n = 5950$. Then $n \notin \varphi(\mathbb{N})$.
- $n = \underbrace{55\ldots5}_{\ell}0$ with $\ell \geq 3$.

Therefore, it remains to show that if $n = \underbrace{55\ldots5}_{\ell}0$ with $3 \leq \ell \leq 10$, then $n \notin \varphi(\mathbb{N})$. To show this, assume that $\varphi(m) = n$ for some $m \in \mathbb{N}$. Since $4 \nmid n$, we have $4 \nmid m$ and m has exactly one odd prime divisor, i.e., $m = p^k$ or $m = 2p^k$ with $k \in \mathbb{N}$ and odd $p \in \mathbb{P}$. In both cases we must have $p^{k-1}(p-1) = n$. Using a computer algebra system, we find the prime factorization of each of the values of n under consideration, namely,

$$\begin{aligned} 5550 &= 2 \cdot 3 \cdot 5^2 \cdot 37, & 55550 &= 2 \cdot 5^2 \cdot 11 \cdot 101, \\ 555550 &= 2 \cdot 5^2 \cdot 41 \cdot 271, & 5555550 &= 2 \cdot 3 \cdot 5^2 \cdot 7 \cdot 11 \cdot 13 \cdot 37, \\ 55555550 &= 2 \cdot 5^2 \cdot 239 \cdot 4649, & 555555550 &= 2 \cdot 5^2 \cdot 11 \cdot 73 \cdot 101 \cdot 137, \\ 5555555550 &= 2 \cdot 3^2 \cdot 5^2 \cdot 37 \cdot 333667, & 55555555550 &= 2 \cdot 5^2 \cdot 11 \cdot 41 \cdot 271 \cdot 9091. \end{aligned}$$

Thus if $p^{k-1} \mid n$ with $k \geq 3$, then $k = 3$ and $p = 3$ or 5, and so $p^{k-1}(p-1) \neq n$. Further, it is easy to check that there is no prime number p with $p(p-1) = n$. Hence $k = 1$ and $n+1$ is a prime. However, all the numbers $\underbrace{55\ldots5}_{\ell}1, 3 \leq \ell \leq 10$, are composite, and this gives the desired contradiction. q.e.d.

It is not difficult to find minimal sets for some sets of shifted values of the Euler totient function. For example, it can be proved that (this result is due to I. Baoulina)

$$\mathcal{S}(3+\varphi(\mathbb{N})) = \{4,5,7,9,11,13,21,23,31,33,61,63,81,83\},$$

where $3+\varphi(\mathbb{N}) := \{3+\varphi(n) : n \in \mathbb{N}\}$. We leave this as an exercise for the interested reader.

Related to Euler's totient function but less well-known is the Dedekind ψ-function, defined by

$$\psi(n) := \prod_{p|n}(p+1)p^{v_p-1} \text{ if } n = \prod_{p|n} p^{v_p}, \quad \psi(1) := 1.$$

Again $\psi(n)$ is even for $n \geq 3$. As in the proof of the corresponing theorem for the Euler totient function, we need to prove that some natural numbers cannot occur as values of the ψ-function. This can be done analogously and will thus be omitted here. As we shall see in the proof of the following theorem, the determination of the corresponding minimal set is possible, although it requires a little more effort than in the case of the Euler totient function.

Theorem II.1.7 *Let ψ be the Dedekind ψ-function and $\psi(\mathbb{N}) := \{\psi(m) : m \in \mathbb{N}\}$. Then*

$$\mathcal{S}(\psi(\mathbb{N})) = \{1,3,4,6,8,20,72,90,222,252,500,522,552,570,592,750,770,\\ 992,7000,\underbrace{55\ldots5}_{69}0\}.$$

Proof. First note, that the elements given above are pairwise incomparable and we have

$$\begin{array}{rclrclrclrcl}
\psi(1) &=& 1, & \psi(2) &=& 3, & \psi(3) &=& 4, & \psi(5) &=& 6,\\
\psi(7) &=& 8, & \psi(19) &=& 20, & \psi(71) &=& 72, & \psi(89) &=& 90,\\
\psi(146) &=& 222, & \psi(251) &=& 252, & \psi(499) &=& 500, & \psi(521) &=& 522,\\
\psi(411) &=& 552, & \psi(569) &=& 570, & \psi(511) &=& 592, & \psi(625) &=& 750,\\
\psi(769) &=& 770, & \psi(991) &=& 992, & \psi(6631) &=& 7000, & & &
\end{array}$$

and

$$\psi(\underbrace{55\ldots5}_{68}49) = \underbrace{55\ldots5}_{69}0,$$

since $\underbrace{55\ldots5}_{68}49$ is a prime, so all these numbers belong to $\psi(\mathbb{N})$.

Let $n \in \psi(\mathbb{N})$ be arbitrary with at least two digits. If $d \triangleleft n$ with $d \in \{1,3,4,6,8\}$, then $n \notin \mathcal{S}(\psi(\mathbb{N}))$. So n only contains the digits $2,5,7,9,0$. First we note that $22,50,52,70,92,292,502,550,700,922,952 \notin \psi(\mathbb{N})$. For all other natural numbers with exactly two or three digits that contain only the digits $2,5,7,9,0$, we find by distinguishing cases that they either belong to $\mathcal{S}(\psi(\mathbb{N}))$ or can be generated by an element of $\mathcal{S}(\psi(\mathbb{N}))$. Now suppose n has at least four digits. Then one of the following cases holds:

- n ends with a 2 and contains a 7. Then $72 \triangleleft n$.
- n ends with a 2 and contains a zero but no 7. Then $20 \triangleleft n$ or $90 \triangleleft n$ or $502 \triangleleft n$.
- n ends with a 2 and contains only the digits 2 and 5. Then $222 \triangleleft n$ or $552 \triangleleft n$.
- n ends with a 2 and contains only the digits 2 and 9. Then $222 \triangleleft n$ or $992 \triangleleft n$.
- n ends with a 2 and contains both the digits 5 and 9. Then $592 \triangleleft n$ or $952 \triangleleft n$.
- n ends with a zero and contains a 2 or 9. Then $20 \triangleleft n$ or $90 \triangleleft n$.
- n ends with a zero and contains a 7 and a 5. Then $570 \triangleleft n$ or $750 \triangleleft n$.
- n ends with a zero and contains only the digits 7 and 0. Then $770 \triangleleft n$ or $700 \triangleleft n$.
- n ends with a zero and contains only the digits 5 and 0. Then $550 \triangleleft n$ or $500 \triangleleft n$.

Thus, there are only four cases in which possibly $n \in \mathcal{S}(\psi(\mathbb{N}))$, namely $502 \triangleleft n$, $550 \triangleleft n$, $700 \triangleleft n$, $952 \triangleleft n$. If $952 \triangleleft n$ and n ends with a 2, we only need to consider those n with $7 \not\triangleleft n$. Then n contains one of the following strings:

$$9952, 9592, 5952, 9552, 2952, 9252, 9522, 9052, 9502.$$

In each of these cases there is an $x \in \mathcal{S}(\psi(\mathbb{N}))$ with $x < n$, $x \triangleleft n$, so $n \notin \mathcal{S}(\psi(\mathbb{N}))$. If $700 \triangleleft n$ and n ends with a zero, we have one of the following cases:

- n contains a 2. Then $20 \triangleleft n$.
- n contains a 9. Then $90 \triangleleft n$.
- n contains another 7. Then $770 \triangleleft n$.

- n contains a 5. Then $750 \triangleleft n$ or $570 \triangleleft n$.
- $n \in 7\{0\}^*$. Then $7000 \triangleleft n$.

If $502 \triangleleft n$ and n ends with a 2, again we only need to consider those n with $7 \not\triangleleft n$. Then n contains one of the following strings:

$$5002, 2502, 5202, 5022, 5502, 5052, 9502, 5902, 5092.$$

In each of these cases there is an $x \in \mathcal{S}(\psi(\mathbb{N}))$ with $x < n$, $x \triangleleft n$, so $n \notin \mathcal{S}(\psi(\mathbb{N}))$. If $550 \triangleleft n$, the only numbers n that need to be examined are of the form $55\ldots50$, since otherwise $500 \triangleleft n$. It suffices to show that if $n = \underbrace{55\ldots5}_{\ell}0$ with $3 \leq \ell \leq 68$, then there is no $m \in \mathbb{N}$ such that $\psi(m) = n$.

We note that if m has k distinct odd prime factors, then $2^k \mid \psi(m)$. In our case, $n \equiv 2 \bmod 4$, so if there is an $m \in \mathbb{N}$ with $\psi(m) = n$, then $m = p^k$ or $m = 2p^k$ with $p \in \mathbb{P}$ odd. Hence either $p^{k-1}(p+1) = n$ or $3p^{k-1}(p+1) = n$ with $k \in \mathbb{N}$ and odd $p \in \mathbb{P}$. Again a computer search shows that if $p^{k-1} \mid n$ with $k \geq 3$, then $(k,p) \in \{(3,3),(3,5),(3,7),(3,11),(4,3)\}$, which implies that $p^{k-1}(p+1) < 3p^{k-1}(p+1) \leq 3 \cdot 11^2 \cdot 12 < 5550 \leq n$. Next, one can easily verify that there is no prime p with $n = p(p+1)$ or $n = 3p(p+1)$, and so $k \neq 2$. Finally, note that neither $n-1$ nor $(n/3)-1$ is a prime, which yields $k \neq 1$. This concludes the proof. q.e.d.

II.1.4 Congruence classes

In this section we determine the minimal sets of some congruence classes. Since the numbers in a given congruence class are "well-distributed" it is relatively easy to determine the minimal sets, at least for $m \leq 6$. In fact, one even gets an algorithm for determining all minimal sets to a given modulus.

We begin with the modulus 2. This result is immediate.

Theorem II.1.8 *We have*

$$\mathcal{S}([0]_2) = \{2,4,6,8,10,30,50,70,90\},$$
$$\mathcal{S}([1]_2) = \{1,3,5,7,9\}.$$

The cases $m = 3$ and $m = 4$ require a little more work.

Theorem II.1.9 *We have*

$$\begin{aligned}\mathcal{S}([0]_3) = \{&3,6,9,12,15,18,21,24,27,42,45,48,51,54,57,72,75,78,81,84,87,\\ &111,114,117,141,144,147,171,174,177,222,225,228,252,255,258,\\ &282,285,288,411,414,417,441,444,447,471,474,477,522,525,528,\\ &552,555,558,582,585,588,711,714,717,741,744,747,771,774,777,\\ &822,825,828,852,855,858,882,885,888\},\\ \mathcal{S}([1]_3) = \{&1,4,7,22,25,28,52,55,58,82,85,88\},\\ \mathcal{S}([2]_3) = \{&2,5,8,11,14,17,41,44,47,71,74,77\}.\end{aligned}$$

Proof.

- We begin with the minimal set of $[1]_3$. Let $x \in [1]_3$. If x contains a digit that is congruent 1 modulo 3, we are done. If x contains a digit d that is congruent 0 modulo 3, we can remove d from x. The resulting number y will be congruent 1 modulo 3 and we have $y \lhd x$. Hence no element of $\mathcal{S}([1]_3)$ can contain a digit that is congruent 0 modulo 3. It remains to consider numbers all of whose digits are congruent 2 modulo 3. But then $x \in \{2,5,8\}^*$. Therefore the result follows.

- Determining the minimal set of $[2]_3$ is completely analogous to that of $[1]_3$.

- Now we determine the minimal set of $[0]_3$. Let $x \in [0]_3$. If $3 \lhd x$, $6 \lhd x$, or $9 \lhd x$, we are done. If $0 \lhd x$, then the number y obtained by removing 0 is congruent 0 modulo 3 and we have $y \lhd x$. Thus it remains to consider numbers all of whose digits are not congruent 0 modulo 3. Suppose that x has exactly two digits. Then one of these digits has to be congruent 1 modulo 3 and the other one has to be congruent 2 modulo 3. All these numbers belong to $\mathcal{S}([0]_3)$. Now suppose that x has exactly 3 digits. Then either all of these digits are congruent 1 modulo 3 or all digits are congruent 2 modulo 3. Again all those numbers belong to $\mathcal{S}([0]_3)$. Finally, let x have at least 4 digits. If x has three digits that are congruent 1 modulo 3 or three digits that are congruent 2 modulo 3, these digits form a number y with $y \in \mathcal{S}([0]_3)$ and $y \lhd x$. If this is not the case, x has at least one digit that is congruent 1 modulo 3 and one digit that is congruent 2 modulo 3. Again these digits form a number y with $y \in \mathcal{S}([0]_3)$ and $y \lhd x$. In either case $x \notin \mathcal{S}([0]_3)$ and this completes the proof.

q.e.d.

Theorem II.1.10 *We have*

$$\mathcal{S}([0]_4) = \{4, 8, 12, 16, 20, 32, 36, 52, 56, 60, 72, 76, 92, 96, 100, 300, 500, 700, 900\},$$
$$\mathcal{S}([1]_4) = \{1, 5, 9, 33, 37, 73, 77\},$$
$$\mathcal{S}([2]_4) = \{2, 6, 10, 14, 18, 30, 34, 38, 50, 54, 58, 70, 74, 78, 90, 94, 98\},$$
$$\mathcal{S}([3]_4) = \{3, 7, 11, 15, 19, 51, 55, 59, 91, 95, 99\}.$$

Proof. We will use a general observation about congruence classes modulo 4 in this proof: If $x \in [a]_4$ has at least three digits, then the number y that is obtained from x by taking only the last two digits is also in $[a]_4$. Thus if $y \neq 0$, we have $x \notin \mathcal{S}([a]_4)$. Hence for $a \in \{1, 2, 3\}$, every element of $\mathcal{S}([a]_4)$ has at most two digits.

- Let $x \in [1]_4$. If x has exactly one digit, then $x \in \{1, 5, 9\}$. If x has exactly two digits, then both digits have to be congruent 3 modulo 4. Hence $x \in \{3, 7\}^*$ and we are done.
- Let $x \in [2]_4$. If x has exactly one digit, then $x \in \{2, 6\}$. If x has exactly two digits, then the first digit is congruent 1 or 3 modulo 4 and the second digit is congruent 0 modulo 4.
- Determining the minimal set of $[3]_4$ is completely analogous to that of $[1]_4$.
- Determining the minimal set of $[0]_4$ is nearly the same as for $[2]_4$ with one difference: If x has at least three digits, then the number obtained by taking only the last two digits is congruent 0 modulo 4, but it is only contained in $\mathcal{S}([0]_4)$ if it is not 0. So we have to add to $\mathcal{S}([0]_4)$ all the numbers that have three digits, end with 00, and whose first digit is congruent 1 or 3 modulo 4.

q.e.d.

The minimal sets for $m = 5$ are as easy to determine as those for $m = 2$.

Theorem II.1.11 *We have*

$$\mathcal{S}([0]_5) = \{5, 10, 20, 30, 40, 60, 70, 80, 90\},$$
$$\mathcal{S}([1]_5) = \{1, 6\},$$
$$\mathcal{S}([2]_5) = \{2, 7\},$$
$$\mathcal{S}([3]_5) = \{3, 8\},$$
$$\mathcal{S}([4]_5) = \{4, 9\}.$$

Let us now consider the case $m = 6$.

Theorem II.1.12 *We have*

$$\begin{aligned}
\mathcal{S}([0]_6) = \{&6, 12, 18, 24, 30, 42, 48, 54, 72, 78, 84, 90, 114, 144, 150, 174, 210, 222,\\
&228, 252, 258, 270, 282, 288, 414, 444, 450, 474, 510, 522, 528, 552, 558,\\
&570, 582, 588, 714, 744, 750, 774, 810, 822, 828, 852, 858, 870, 882, 888,\\
&1110, 1170, 1410, 1470, 1710, 1770, 2250, 2550, 2850, 4110, 4170, 4410,\\
&4470, 4710, 4770, 5250, 5550, 5850, 7110, 7170, 7410, 7470, 7710, 7770,\\
&8250, 8550, 8850\},\\
\mathcal{S}([1]_6) = \{&1, 7, 25, 43, 49, 55, 85, 223, 229, 283, 289, 445, 523, 529, 583, 589, 823,\\
&829, 883, 889\},\\
\mathcal{S}([2]_6) = \{&2, 8, 14, 44, 50, 56, 74, 110, 116, 170, 176, 410, 416, 470, 476, 554, 710,\\
&716, 770, 776\},\\
\mathcal{S}([3]_6) = \{&3, 9, 15, 21, 27, 45, 51, 57, 75, 81, 87, 111, 117, 141, 147, 171, 177, 225,\\
&255, 285, 411, 417, 441, 447, 471, 477, 525, 555, 585, 711, 717, 741, 747,\\
&771, 777, 825, 855, 885\},\\
\mathcal{S}([4]_6) = \{&4, 10, 16, 22, 28, 52, 58, 70, 76, 82, 88, 112, 118, 172, 178, 250, 256, 550,\\
&556, 712, 718, 772, 778, 850, 856\},\\
\mathcal{S}([5]_6) = \{&5, 11, 17, 23, 29, 41, 47, 71, 77, 83, 89, 143, 149, 221, 227, 281, 287, 443,\\
&449, 743, 749, 821, 827, 881, 887\}.
\end{aligned}$$

Proof. We will sketch the proof for the first case. The others follow similarly.

If $x \in [0]_6$, its last digit can only be one of $0, 2, 4, 6, 8$. If its last digit is 6, we are done. If its last digit is 2 or 8, the number obtained by removing the last digit is congruent 1 modulo 3. So we just have to check the list of $\mathcal{S}([1]_3)$ in Theorem II.1.9 and append to each of these numbers the digit 2 or 8. If the last digit of x is 4, the process is analogous, here we have to take the list of $\mathcal{S}([2]_3)$. If the last digit of x is 0, the number obtained by removing the last digit is congruent 0 modulo 3. If it is also congruent 0 modulo 6, then there already is an $z \in \mathcal{S}([0]_6)$ with $z \lhd x$. So we have to take all the elements of $\mathcal{S}([0]_3)$ that are not congruent 0 modulo 6 and append the digit 0. q.e.d.

It is equally easy to determine the minimal set of M if M is a union of congruence classes.

Theorem II.1.13 *Let*

$$\square \bmod 6 := \{n \in \mathbb{N} : n \equiv x^2 \bmod 6 \textit{ for some } x \in \mathbb{N}\}.$$

Then

$$\mathcal{S}(\square \bmod 6) = \{1, 3, 4, 6, 7, 9, 22, 25, 28, 52, 55, 58, 82, 85, 88\}.$$

Proof. It is easy to check that $n \in \square \bmod 6$ if and only if $n \equiv 1, 3, 4, 6 \bmod 6$. So we have indeed $\{1, 3, 4, 6, 7, 9, 22, 25, 28, 52, 55, 58, 82, 85, 88\} \subset \square \bmod 6$. Moreover, these elements are pairwise incomparable. Let now $n \in \square \bmod 6$ be arbitrary. If $d \triangleleft n$ with $d \in \{1, 3, 4, 6, 7, 9\}$, we are done. So suppose that n contains none of these digits. Assume first that n contains exactly two digits. Then $22, 25, 28, 52, 55, 58, 82, 85, 88 \in \square \bmod 6$ and $20, 50, 80 \notin \square \bmod 6$. If n has more than two digits, then either two of its digits are nonzero and hence there is an $x \in \mathcal{S}(\square \bmod 6)$ with $x < n$ and $x \triangleleft n$, or n has only one nonzero digit. In the latter case, since this digit is congruent 2 modulo 3, we have $n \equiv 2 \bmod 6$, and therefore $n \notin \square \bmod 6$. q.e.d.

Theorem II.1.14 *Let*

$$\square \bmod 7 := \{n \in \mathbb{N} : n \equiv x^2 \bmod 7 \textit{ for some } x \in \mathbb{N}\}.$$

Then

$$\mathcal{S}(\square \bmod 7) = \{1, 2, 4, 7, 8, 9, 30, 35, 36, 50, 53, 56, 60, 63, 65, 333, 555, 666\}.$$

Proof. Since $n \in \square \bmod 7$ if and only if $n \equiv 1, 2, 4, 7 \bmod 7$, we have

$$\{1, 2, 4, 7, 8, 9, 30, 35, 36, 50, 53, 56, 60, 63, 65, 333, 555, 666\} \subset \square \bmod 7,$$

and again these elements are pairwise incomparable. Let $n \in \square \bmod 7$. If $d \triangleleft n$ with $d \in \{1, 2, 4, 7, 8, 9\}$, we are done. So suppose that n contains none of these digits. If n has exactly two digits, then $30, 35, 36, 50, 53, 56, 60, 63, 65 \in \square \bmod 7$ and $33, 55, 66 \notin \square \bmod 7$. If n has at least three digits and at least two of them are distinct, we are done. It therefore remains to assume that n is a d-repdigit with $d \in \{3, 5, 6\}$. Then either $333 \triangleleft n$ or $555 \triangleleft n$ or $666 \triangleleft n$. q.e.d.

In the above cases it was relatively easy to get the minimal sets (in fact we did not always construct them but this is possible with the above ideas). If we try the same approach for the congruence classes modulo 7, we see that it gets

more complicated. Suppose that we want to get the minimal set of $[0]_7$. If the last digit of $x \in [0]_7$ is 0, then the number obtained by striking away this digit will be congruent 2 modulo 7. So we would have to look at the minimal set of $[2]_7$, which we do not know yet. Thus, in general, one cannot construct the minimal set of a congruence class on its own. But it is possible to construct the minimal sets of all m congruences classes modulo m simultaneously with Algorithm II.1.1.

Algorithm II.1.1 Minimal set algorithm for congruence classes

1: **for all** $a \in \{0, \ldots, m-1\}$ **do**
2: Initialize $\mathcal{S}([a]_m)$ as empty list
3: **end for**
4: **for all** $a \in \{0, \ldots, m-1\}$ and each digit $d \in \{1, \ldots, 9\}$ **do**
5: Determine which digit d lies in which congruence class $[a]_m$
6: Write these digits in $\mathcal{S}([a]_m)$
7: **end for**
8: **for all** $a \in \{0, \ldots, m-1\}, b \in \{0, \ldots, 9\}$ **do**
9: Determine all $x \in \mathbb{Z}/m\mathbb{Z}$ (depending on a and b) with $10x + b \equiv a \bmod m$
10: Let $X(a, b)$ be the set of all such x
11: **end for**
12: **for all** $a \in \{0, \ldots, m-1\}, b \in \{0, \ldots, 9\}, x \in X(a, b)$ **do**
13: Take every $y \in \mathcal{S}([x]_m)$ and test if there is a number $z \in \mathcal{S}([a]_m)$ such that $z \lhd y * b$
14: **if** there is no such z **then**
15: Add $y * b$ to $\mathcal{S}([a]_m)$
16: **end if**
17: **end for**
18: **repeat**
19: Lines 12 to 17
20: **until** no number gets added in any minimal set

Examples of minimal sets determined with this algorithm can be found in Appendix A.3.

Algorithm II.1.1 is just a special case of Algorithm II.1.2. This algorithm can be applied to specific partitions of $\mathbb{N}$.

Definition II.1.15 Let $m \in \mathbb{N}$ and $A_1, \ldots, A_m$ be a partition of $\mathbb{N}$. The partition is called **truncating stable** if it has the following property:

For any $i \in \{1, \ldots, m\}$ and any digit $b \in \{0, \ldots, 9\}$ there is a set $J_i(b) \subset \{1, \ldots, m\}$ such that $10a + b \in A_i$ if and only if there is a $j \in J_i(b)$ with $a \in A_j$.

This condition should be read as follows: Let $n \in A_i$. Then for any digit d we can predict in which of the sets A_l the number $n * d$ lies (note that for distinct

i_1, i_2 the sets $J_{i_1}(b)$ and $J_{i_2}(b)$ are disjoint since A_{i_1} and A_{i_2} are disjoint). If this condition is fulfilled, we can use Algorithm II.1.2 to determine the minimal sets of the sets A_i simultaneously. The condition is, for example, fulfilled for (unions of) congruence classes, but not fulfilled if we partition the natural numbers in primes and composite numbers.

Algorithm II.1.2 Minimal set algorithm for truncating stable partitions

1: **for all** $i \in \{1, \ldots, m\}$ **do**
2: Initialize $\mathcal{S}(A_i)$ as empty list
3: **end for**
4: **for all** $i \in \{1, \ldots, m\}$ and each digit $d \in \{1, \ldots, 9\}$ **do**
5: Determine which digit d lies in which set A_i
6: Write these digits in $\mathcal{S}(A_i)$
7: **end for**
8: **for all** $i \in \{1, \ldots, m\}, b \in \{0, \ldots, 9\}, j \in J_i(b)$ **do**
9: Take every $y \in \mathcal{S}(A_j)$ and test if there is a number $z \in \mathcal{S}(A_i)$ such that $z \triangleleft y * b$
10: **if** there is no such z **then**
11: Add $y * b$ to $\mathcal{S}(A_i)$
12: **end if**
13: **end for**
14: **repeat**
15: Lines 8 to 13
16: **until** no number gets added in any minimal set

We will now prove the validity of Algorithm II.1.2.

Theorem II.1.16 *Let $m \in \mathbb{N}$ and $A_1, \ldots, A_m$ be a truncating stable partition of $\mathbb{N}$. Then Algorithm II.1.2 terminates after finite time and gives the minimal sets of the sets A_i.*

Proof. First we show that the algorithm terminates after finite time. Each of the **for** loops needs only finite time (since all involved sets are finite), so we just have to consider the **repeat** statement. The respective steps will only be repeated finitely often, since we know that minimal sets are finite. Therefore the algorithm will terminate after finite time.

Now we show that the algorithm indeed constructs the minimal sets for the sets A_i. It is clear that the constructed sets are not truncatable and that each of them is a subset of the respective set A_i. We have to show that any natural number $n \in A_i$ can be generated by some natural number $\widetilde{n} \in \mathcal{S}(A_i)$. For that, we show that each number n with l digits ($l \in \mathbb{N}$) either lies in one of the constructed sets or there is an $\widetilde{n} \in \mathbb{N}$ that lies in the respective constructed set with $\widetilde{n} \triangleleft n$. Suppose that the algorithm terminates after the **for** loop in line 8 has been gone through

k times. In the first $k-1$ steps all natural numbers n with at most k digits have been considered and for each of them either $n \in \mathcal{S}(A_i)$ for some i or there is an i and an $\widetilde{n} \in \mathbb{N}$ with $n, \widetilde{n} \in A_i, \widetilde{n} \triangleleft n$. So the claim holds for $l \leq k$ and we are left to consider natural numbers with at least $k+1$ digits. We show by induction on $l > k$ that none of them can belong to a minimal set.

When the **for** loop in line 8 is executed the k-th time, all natural numbers with $k+1$ digits have been considered and none of them belongs to a minimal set (otherwise the algorithm would not have terminated). Now suppose that no natural number with l digits (where $l \geq k+1$) is in a minimal set. We show that the same holds for all natural numbers with $l+1$ digits. Let $x \in A_i \subset \mathbb{N}$ with $\#x = l+1$. Write $x = y * b$ with $\#y = l, b \in \{0, \ldots, 9\}$. Then we know that $y \in A_j$ for some $j \in J_i(b)$ and from the induction hypothesis we know that there is a $z \in A_j$ with $\#z < l$ and $z \triangleleft y$. Let $w := z * b$. Then $\#w < l+1$ and $w \triangleleft x$. Since the sets $J_i(b)$ are disjoint for distinct i, we also get $w \in A_i$. This proves the theorem. q.e.d.

Remark II.1.17 Since congruence classes define a truncation stable partition of $\mathbb{N}$ and Algorithm II.1.1 is just a special case of Algorithm II.1.2, we know that Algorithm II.1.1 terminates after a finite number of steps and gives the minimal sets of the congruence classes modulo m.

Although we know that Algorithm II.1.2 will terminate, we do not know when, hence we cannot determine the run-time of the algorithm in general. However, for congruence classes we can say something about the maximal number of digits in the minimal sets.

Theorem II.1.18 *Let* $m = 2^a 5^b$.

1. *The largest number in* $\mathcal{S}([0]_m)$ *has exactly* $\max(a,b)+1$ *digits.*
2. *For all* $k \not\equiv 0$, *the largest number in* $\mathcal{S}([k]_m)$ *has at most* $\max(a,b)$ *digits.*
3. *There is a* $k \not\equiv 0$ *such that the largest number in* $\mathcal{S}([k]_m)$ *has exactly* $\max(a,b)$ *digits.*

Proof. Since $m = 2^a 5^b$, the congruence class modulo m of $n \in \mathbb{N}$ depends only on its last $\max(a,b)$ digits. If n has at least $\max(a,b)+1$ digits, the number formed by its last $\max(a,b)$ digits is in the same congruence class. Hence n cannot be in $\mathcal{S}([k]_m)$ except for the case when the last $\max(a,b)$ digits are all 0. This can only happen if $n \equiv 0 \bmod m$, i.e., if $k = 0$. If $n \equiv 0 \bmod m$ and n has at least $\max(a,b)+2$ digits, either one of the last $\max(a,b)$ digits is not zero (then n cannot be in $\mathcal{S}([0]_m)$ due to the arguments above) or the number formed by the first

digit and the last $\max(a,b)$ digits is nonzero and congruent 0 modulo m (since all digits except for the first are 0). In either case, n cannot be in the minimal set.

It remains to show that there is an element in $\mathcal{S}([0]_m)$ with $\max(a,b)+1$ digits and that there is a $k \neq 0$ such that there is an element in $\mathcal{S}([k]_m)$ with $\max(a,b)$ digits.

In the first case, we note that $n = 10^{\max(a,b)}$ has exactly $\max(a,b)+1$ digits and is congruent 0 modulo m. The only numbers $x \in \mathbb{N}$ with $x \lhd n$ and $x \neq n$ are the numbers $x = 10^y$ with $y < \max(a,b)$ and none of these is congruent to 0 modulo m.

For the second case, take $k = m-1$ and $n = 10^{\max(a,b)} - 1$. Then $n \in [k]_m$, n has exactly $\max(a,b)$ digits, and all of its digits are 9. So there is an $x \in \mathcal{S}([k]_m)$ with $x \neq n, x \lhd n$ if and only if there is a $y < \max(a,b)$ with $10^y - 1 \equiv -1 \bmod m$. But this is not possible. q.e.d.

Moreover, Table A.3.2 in Appendix A.3.2 gives rise to the following conjecture:

Conjecture II.1.19 *Let $m \in \mathbb{N}$ with $\gcd(m,10) = 1$.*

1. *For any $k \neq 0$, the largest number in $\mathcal{S}([k]_m)$ has at most $m-1$ digits. If 10 is a primitive root modulo m, then there is a natural number in $\mathcal{S}([k]_m)$ with exactly $m-1$ digits.*

2. *The largest number in $\mathcal{S}([0]_m)$ has at most m digits. If 10 is a primitive root modulo m, then there is a natural number in $\mathcal{S}([0]_m)$ with exactly m digits.*

II.1.5 Basic set operations

The examples from the previous sections already indicate that minimal sets behave rather unexpectedly. It is easy to obtain some minimal sets in explicit form, while for other sets the explicit form of the corresponding minimal set may only be achieved conditionally (compare Section II.1.7). Already the size of minimal sets seems to be almost unpredictable.

For $k \in \mathbb{N}$ define $A_k := \mathbb{N} \cap [10^{k-1}, 10^k)$ (i.e., A_k is the set of natural numbers with k digits in base 10). One easily verifies $\mathcal{S}(A_k) = A_k$ which shows that the minimal set can be as large as we please even for finite sets. Moreover, A_k is minimal among all sets M with $|\mathcal{S}(M)| = 9 \cdot 10^{k-1}$. In general it seems to be difficult to prove an upper bound for the number of minimal elements in an arbitrary set.

We show that almost no structural relationship between sets is carried over to the respective minimal sets. In this section, let $M, L \subset \mathbb{N}$ be infinite sets and $F \subset \mathbb{N}$ be a finite set. We will need some more notation.

For $a, b \in \mathbb{N}$ we set

$$\{a * b\}^| := \{n \in \mathbb{N} : n = a^{*k} * b^{*l} \text{ for some } k, l \in \mathbb{N}_0\}.$$

For example,

$$\{1 * 37\}^| = \{1, 37, 11, 137, 3737, 111, 1137, 13737, 373737, \ldots\}.$$

We begin to study subsets. Let $L \subset M \subset \mathbb{N}$. For arbitrary subsets it seems impossible to deduce properties (consider $L = \mathbb{P}$ and $M = \mathbb{N}$). If $F \cap \mathcal{S}(M) = \emptyset$, then it is obvious that $\mathcal{S}(M\backslash F) = \mathcal{S}(M)$. In the other case, one would expect that $\mathcal{S}(M\backslash F)$ has more elements than $\mathcal{S}(M)$: Removing minimal elements should result in the necessity of larger minimal elements. Since larger minimal elements can generate a smaller fraction of natural numbers, one would expect that one needs more of them. This is not always true.

Theorem II.1.20 *There are infinitely many $M \subset \mathbb{N}$ such that there is a set $F \subset M$ with $|\mathcal{S}(M\backslash F)| < |\mathcal{S}(M)|$.*

Proof. Let $a, b \in \{1, \ldots, 9\}$ and $a \neq b$. Let $\widetilde{M} = \{a * b\}^|$ and $F = \{a, b\}$. Then $\widetilde{M}$ has infinitely many infinite subsets M with $F \subset \{a, b, a * b\} \subset M$ and such that any $x \in M$ with $x \neq a, b$ has two distinct digits. For all these sets we have $\mathcal{S}(M) = F = \{a, b\}$ and $\mathcal{S}(M\backslash F) = \{a * b\}$. q.e.d.

Example II.1.21 Let $\widetilde{M} = \{1 * 6\}^|$, i.e., $\widetilde{M}$ consists exactly of the natural numbers, all of whose first digits are 1 and all of whose last digits are 6. Take $F = \{1, 6\}$ and let M be any subset of $\widetilde{M}$ with $\{1, 6, 16\} \in M$ and $\{11, 66, 111, 666, \ldots\} \notin M$. Then $\mathcal{S}(M) = \{1, 6\}$ and $\mathcal{S}(M\backslash F) = \{16\}$.

For "generic" sets (i.e., sets that are not specifically constructed for this purpose) the above phenomenon seems not to happen. This leads to the following definition and conjecture.

Definition II.1.22 For a set $M \subset \mathbb{N}$ define a sequence $\delta^n(M)$ of sets recursively by

$$\delta^0(M) := M, \quad \delta(M) := \delta^1(M) := M\backslash\mathcal{S}(M), \quad \delta^{n+1}(M) := \delta(\delta^n(M))$$

and let $\eta^n(M) := |\mathcal{S}(\delta^n(M))|$ and $\eta(M) := \eta^1(M)$.

Example II.1.23 Let $M := \{1, 6, 16, 1166, 111666, 11116666, \ldots\}$. Then

$$\begin{aligned}
\delta^0(M) &= M,\\
\delta^1(M) &= \{16, 1166, 111666, 11116666, \ldots\},\\
\delta^2(M) &= \{1166, 111666, 11116666, \ldots\},\\
&\vdots
\end{aligned}$$

and

$$\eta^0(M) = 2, \qquad \eta^k(M) = 1 \text{ for all } k \geq 1.$$

As shown in Theorem II.1.20, there are infinitely many M with $\eta(M) < \eta^0(M)$. We conjecture that the number of such sets is "small".

Conjecture II.1.24 *There are only countably many sets $M \subset \mathbb{N}$ with $\eta(M) \leq \eta^0(M)$. For all other sets we have $\eta^n(M) \to \infty$.*

For some of the following results we need to know that for each $k \in \mathbb{N}$ there is an infinite set M with $|\mathcal{S}(M)| = k$.

Lemma II.1.25 *Let $k \in \mathbb{N}$. There are infinitely many sets $M \subset \mathbb{N}$ with $|\mathcal{S}(M)| = k$. In particular there is no bound for the cardinality of minimal sets.*

Proof. Choose m such that there are at least k natural numbers that have exactly m digits. Take k of these numbers and let $\widetilde{M}$ be the finite set that contains these k numbers. Let $\widehat{M} = \langle \widetilde{M} \rangle$. Then $\widehat{M}$ has infinitely many subsets M with $\widetilde{M} \subset M$, i.e., $\mathcal{S}(M) = \widetilde{M}$. q.e.d.

The following theorem can be seen as a generalization of Theorem II.1.20.

Theorem II.1.26 *Let $c, k \in \mathbb{N}$. Then there are infinitely many $M \subset \mathbb{N}$ such that there is an $F \subset M$ with $|\mathcal{S}(M)| = k, |\mathcal{S}(M \backslash F)| = c$.*

Proof. Similar to the proof of Lemma II.1.25 we choose a set $\widetilde{M}$ that has exactly k elements with exactly m digits (for a suitable m) and no other elements.

- If $c \leq k$, take a subset $\widetilde{F} \subset \widetilde{M}$ such that $\left|\widetilde{M} \backslash \widetilde{F}\right| = c$ and take M such that

  $$\widetilde{M} \subset M \subset \langle \widetilde{M} \backslash \widetilde{F} \rangle \cup \widetilde{F}.$$

  Let $F := \widetilde{F}$. Then $\mathcal{S}(M) = \mathcal{S}(\widetilde{M}) = \widetilde{M}$ and $\mathcal{S}(M \backslash F) = \widetilde{M} \backslash F$.

- Now let $c > k$. First we note that the number of elements $x \in \langle \widetilde{M} \rangle$ with $\#x = n$ is strictly increasing as n gets bigger. We choose n such that $n > m$

and $\langle\widetilde{M}\rangle$ has at least c elements with exactly n digits. Let $F \subset \langle\widetilde{M}\rangle$ be the subset that contains all the elements in $\langle\widetilde{M}\rangle$ with less than n digits and choose $M \subset \langle\widetilde{M}\rangle$ such that $F \subset M$, M has exactly c elements with exactly n digits, and all elements of M with more than n digits can be generated by an element of M with exactly n digits. Then we have $\mathcal{S}(M) = \widetilde{M}$, hence $|\mathcal{S}(M)| = k$. Further $\mathcal{S}(M\backslash F)$ contains exactly the c numbers of M that have exactly n digits, thus $|\mathcal{S}(M\backslash F)| = c$.

q.e.d.

Example II.1.27 Let $k = 17$. Then we can take $m = 2$.

1. Let first $c = 13$. We take $\widetilde{M} := \{11, \ldots, 27\}$ and $\widetilde{F} := \{24, \ldots, 27\}$. If we now take M with

$$\{11, \ldots, 27\} \subset M \subset \langle\{11, \ldots, 23\}\rangle \cup \{24, \ldots, 27\},$$

   we get $\mathcal{S}(M) = \{11, \ldots, 27\}$ and $\mathcal{S}(M\backslash F) = \{11, \ldots, 23\}$.

2. Let now $c = 27$ and take again $\widetilde{M} = \{11, \ldots, 27\}$. We choose $n = 3$ and let

$$F = \{11, \ldots, 27\} \text{ and } M = \{11, \ldots, 27\} \cup \langle\{111, \ldots, 137\}\rangle.$$

   Then we get $\mathcal{S}(M) = \{11, \ldots, 27\}$ and $\mathcal{S}(M\backslash F) = \{111, \ldots, 137\}$.

Let us now consider basic set operations. From the known formulae in set theory we can relate some of these operations. Hence we will only consider intersection, union, and complement. Since the intersection of two sets is in particular a subset of each of the sets, the previous theorem applies. Even more unfortunate, the set $\mathcal{S}(M \cap L)$ can be disjoint to $\mathcal{S}(M)$ and $\mathcal{S}(L)$:

Theorem II.1.28 *There are infinitely many sets $L, M \subset \mathbb{N}$ such that*

$$(\mathcal{S}(M) \cup \mathcal{S}(L)) \cap \mathcal{S}(M \cap L) = \emptyset.$$

Proof. Choose L and M such that $\mathcal{S}(M) \cap L = \emptyset, \mathcal{S}(L) \cap M = \emptyset$, and $M \cap L \neq \emptyset$ (for example, let $L \subset \langle\{1, \ldots, 4\}\rangle$ and $M \subset \langle\{6, \ldots, 9\}\rangle$). Then $(\mathcal{S}(M) \cup \mathcal{S}(L))$ and $\mathcal{S}(M \cap L)$ are disjoint. q.e.d.

The next theorem shows that for the union of two sets, there is at least a little bit of structure.

Theorem II.1.29 *We have $\mathcal{S}(M \cup L) \subset \mathcal{S}(M) \cup \mathcal{S}(L)$.*

Proof. Let $x \in \mathcal{S}(M \cup L)$ and without loss of generality $x \in M$. Suppose that $x \notin \mathcal{S}(M)$. Then there is a $z \in M$ with $z \neq x, z \lhd x$. But since z lies also in $M \cup L$ we have $x \notin \mathcal{S}(M \cup L)$, which is a contradiction. q.e.d.

The left-hand side of the equation in Theorem II.1.29 can be equal to the right-hand side (if, for example, $M = [2]_{10}$ and $L = [3]_{10}$) and smaller (if, for example, $M = \{2\} \cup \{p \in \mathbb{P} : p \equiv 1 \bmod 4\}$ and $L = \{p \in \mathbb{P} : p \equiv 3 \bmod 4\}$). In general, equality cannot hold, since the set $\mathcal{S}(M) \cup \mathcal{S}(L)$ could be truncatable. If it is not truncatable, we do indeed have equality:

Theorem II.1.30 *We have $\mathcal{S}(M \cup L) = \mathcal{S}(M) \cup \mathcal{S}(L)$ if and only if $\mathcal{S}(M) \cup \mathcal{S}(L)$ is not truncatable.*

Proof. We know

$$\begin{aligned}\mathcal{S}(M) \subset \mathcal{S}(M \cup L) &\Leftrightarrow \forall x \in \mathcal{S}(M) : x \in \mathcal{S}(M \cup L) \\ &\Leftrightarrow \forall x \in \mathcal{S}(M) : \nexists y \in M \cup L \text{ such that } y \lhd x, y \neq x \\ &\Leftrightarrow \forall x \in \mathcal{S}(M) : \nexists y \in L \text{ such that } y \lhd x, y \neq x \\ &\Leftrightarrow \forall x \in \mathcal{S}(M) : \nexists y \in \mathcal{S}(L) \text{ such that } y \lhd x, y \neq x,\end{aligned}$$

so equality holds if and only if

- $\forall x \in \mathcal{S}(M) : \nexists y \in \mathcal{S}(L)$ such that $y \lhd x, y \neq x$ and
- $\forall x \in \mathcal{S}(L) : \nexists y \in \mathcal{S}(M)$ such that $y \lhd x, y \neq x$,

which is equivalent to saying that $\mathcal{S}(M) \cup \mathcal{S}(L)$ is not truncatable. q.e.d.

From Theorem II.1.29 we get that $|\mathcal{S}(M \cup L)| \leq |\mathcal{S}(M)| + |\mathcal{S}(L)|$. The next theorem shows that more cannot be said.

Theorem II.1.31 *Let $k, c_1, c_2 \in \mathbb{N}$ with $k \leq c_1 + c_2$.*

1. *If $k = 1$, there are sets $M, L \subset \mathbb{N}$ with $|\mathcal{S}(M \cup L)| = 1, |\mathcal{S}(M)| = 1$, and $|\mathcal{S}(L)| = c_2$.*
2. *If $k > 1$, there are sets $M, L \subset \mathbb{N}$ with $|\mathcal{S}(M \cup L)| = k, |\mathcal{S}(M)| = c_1$, and $|\mathcal{S}(L)| = c_2$.*

Proof.

1. Let $z \in \mathbb{N}$. We will construct sets M and L with $\mathcal{S}(M \cup L) = \{z\}$. Choose m such that $m > \#z$ and there are at least c_2 natural numbers with exactly $m - \#z$ digits. Let L be an infinite set with the following properties (it is easy to see that such sets do exist).

- $z \notin L$,
- $L \subset \langle\{z\}\rangle$,
- L has exactly c_2 elements with exactly m digits,
- L has no elements with less than m digits,
- all elements of L with more than m digits can be generated by an element of L with exactly m digits,
- $\langle\{z\}\rangle \backslash L$ is infinite.

Let $M = \langle\{z\}\rangle \backslash L$. Then $\mathcal{S}(M \cup L) = \mathcal{S}(M) = \{z\}$ and $|\mathcal{S}(L)| = c_2$.

2. Let $k > 1$ and $z_1, \ldots, z_k \in \mathbb{N}$ such that there is a digit z_0 such that $z_0, z_1, \ldots, z_k$ are pairwise incomparable. We will construct sets M, L with $|\mathcal{S}(M)| = c_1$, $|\mathcal{S}(L)| = c_2$, and $\mathcal{S}(M \cup L) = \{z_1, \ldots, z_k\}$.

   First suppose that $c_i \geq k$ for $i = 1, 2$. For $j \in \{1, \ldots, k\}$ and $m \in \mathbb{N}$ let

   $$X_{j,m} := \{x \in \mathbb{N} : z_j \triangleleft x, \#x = m\}.$$

   Choose m big enough such that there are sets $\widetilde{M}, \widetilde{L}$ with

   $$\begin{aligned} \widetilde{M} \subset &\{x \in X_{2,m} : x \neq z_2, z_i \not\triangleleft x \text{ for } i \in \{1, 3, \ldots, k\}\} \text{ and } \left|\widetilde{M}\right| = c_1 - k + 1, \\ \widetilde{L} \subset &\{x \in X_{1,m} : x \neq z_1, z_i \not\triangleleft x \text{ for } i \in \{2, 3, \ldots, k\}\} \text{ and } \left|\widetilde{L}\right| = c_2 - k + 1. \end{aligned}$$

   (Such sets would not exist if there were no z_0 such that $z_0, z_1, \ldots, z_k$ are pairwise incomparable.) Now let

   $$M := \{z_1, z_3, z_4, \ldots, z_k\} \cup \langle\widetilde{M}\rangle, \qquad L := \{z_2, z_3, \ldots, z_k\} \cup \langle\widetilde{L}\rangle.$$

   Then we get

   $$\begin{aligned} M \cup L &= \{z_1, \ldots, z_k\} \cup \langle\widetilde{M}\rangle \cup \langle\widetilde{L}\rangle, \\ \mathcal{S}(M) &= \{z_1, z_3, z_4, \ldots, z_k\} \cup \widetilde{M} && \Rightarrow |\mathcal{S}(M)| = c_1, \\ \mathcal{S}(L) &= \{z_2, z_3, \ldots, z_k\} \cup \widetilde{L} && \Rightarrow |\mathcal{S}(L)| = c_2, \\ \mathcal{S}(M \cup L) &= \{z_1, \ldots, z_k\} \end{aligned}$$

   and this gives the result.

Now let $k > c_1$. Pick $c_1 + c_2 - k$ (i.e., none if $k = c_1 + c_2$) elements α_i from $\langle\{z_1, \ldots, z_{c_1}\}\rangle \backslash \{z_1, \ldots, z_{c_1}\}$ such that the set $\{\alpha_1, \ldots, \alpha_{c_1+c_2-k}, z_{c_1+1}, \ldots, z_k\}$ is not truncatable (we can just pick the α_i to have the same number of digits and such that they do not contain any of the $z_i, i > c_1$, as a subsequence. This is possible if we choose the number of digits large enough). Let

$$M := \langle\{z_1, \ldots, z_{c_1}\}\rangle \backslash \{\alpha_1, \ldots, \alpha_{c_1+c_2-k}\}$$

and

$$L := \{\alpha_1, \ldots, \alpha_{c_1+c_2-k}\} \cup \langle\{z_{c_1+1}, \ldots, z_k\}\rangle.$$

Then we get

$$\begin{aligned} M \cup L &= \langle\{z_1, \ldots, z_k\}\rangle, \\ \mathcal{S}(M \cup L) &= \{z_1, \ldots, z_k\}, \\ \mathcal{S}(M) &= \{z_1, \ldots, z_{c_1}\}, \\ \mathcal{S}(L) &= \{\alpha_1, \ldots, \alpha_{c_1+c_2-k}, z_{c_1+1}, \ldots, z_k\}. \end{aligned}$$

Thus the theorem follows.

q.e.d.

Note that if $|\mathcal{S}(M \cup L)| = 1$, we necessarily have $|\mathcal{S}(M)| = 1$ or $|\mathcal{S}(L)| = 1$, hence the first part of Theorem II.1.31 is best possible.

Example II.1.32 Let $k = 7$ and $z_i = i$ for $i = 1, \ldots, 7$. These numbers fulfill the condition mentioned in the proof since we can take $z_0 = 8$ or $z_0 = 9$.

1. Let $c_1 = c_2 = 8$, hence $c_i - k + 1 = 2$. Choose $m = 2$ and let

$$\widetilde{M} = \{28, 82\}, \qquad \widetilde{L} = \{18, 19\}.$$

Then we get

$$\mathcal{S}(M) = \{1, 3, 4, 5, 6, 7, 28, 82\}, \qquad \mathcal{S}(L) = \{2, 3, 4, 5, 6, 7, 18, 19\},$$

and

$$\mathcal{S}(M \cup L) = \{1, 2, 3, 4, 5, 6, 7\}.$$

2. Let $c_1 = 3$ and $c_2 = 5$, hence $c_1 + c_2 - k = 1$. Take $\alpha_1 = 11$ and let

$$M = \langle\{1, 2, 3\}\rangle \backslash \{11\}, \qquad L = \{11\} \cup \langle\{4, 5, 6, 7\}\rangle.$$

Then we get

$$\mathcal{S}(M) = \{1,2,3\}, \qquad \mathcal{S}(L) = \{4,5,6,7,11\},$$

and

$$\mathcal{S}(M \cup L) = \{1,2,3,4,5,6,7\}.$$

With respect to the complement there is also very little structure. It is clear that $\mathcal{S}(M^c)$ contains (in addition to elements with more than one digit) exactly the one-digit numbers that are not in $\mathcal{S}(M)$.

One question that arises is the following: Given a finite set M' that is the minimal set of an infinite set M, can one deduce information about the set $\mathcal{S}(M^c)$? Apart from the one-digit numbers, this is not the case:

Theorem II.1.33 *Let M' be a finite, not truncatable set that contains at least one digit $x \in \{1,\dots,9\}$. Then for any $\varepsilon > 0$ there are two infinite sets M_1, M_2 such that $\mathcal{S}(M_1) = \mathcal{S}(M_2) = M'$ and*

$$\frac{\left|\mathcal{S}(M_1^c) \cap \mathcal{S}(M_2^c)\right|}{\left|\mathcal{S}(M_1^c)\right|} < \varepsilon.$$

If further M' contains all digits, M_1 and M_2 can be chosen such that $\mathcal{S}(M_1^c)$ and $\mathcal{S}(M_2^c)$ are disjoint.

Proof. Without loss of generality let $M' = \{1,\dots,d\} \cup \widetilde{M}$ with $\widetilde{M}$ possibly empty and $\widetilde{M} \cap \{1,\dots,9\} = \emptyset$. Let

$$M_1 := \{1,\dots,d\}^{*\{1,\dots,k\} \cup \{m+2n, n \in \mathbb{N}_0\}} \cup \widetilde{M} \cup \{x \in \{1,\dots,d,0\}^* : 0 \lhd x\},$$
$$M_2 := \{1,\dots,d\}^{*\{1,\dots,m\} \cup \{m+2n, n \in \mathbb{N}_0\}} \cup \widetilde{M} \cup \{x \in \{1,\dots,d,0\}^* : 0 \lhd x\}$$

with some $m, k \in \mathbb{N}$ such that $m > k+1$. Since M' is not truncatable, no number from $\widetilde{M}$ contains one of the digits $1,\dots,d$, thus we have

$$\mathcal{S}(M_1) = \mathcal{S}(M_2) = \{1,\dots,d\} \cup \widetilde{M} = M'.$$

We get

$$M_1^c = \left(\langle\{d+1,\dots,9\}\rangle \cup \{1,\dots,d\}^{*\{k+1,\dots,m-1\} \cup \{m+2n+1, n \in \mathbb{N}_0\}}\right) \setminus \widetilde{M},$$
$$M_2^c = \left(\langle\{d+1,\dots,9\}\rangle \cup \{1,\dots,d\}^{*\{m+2n+1, n \in \mathbb{N}_0\}}\right) \setminus \widetilde{M},$$

and since $\widetilde{M} \cap \{1, \ldots, 9\} = \emptyset$ and no element of $\widetilde{M}$ contains one of the digits $1, \ldots, d$, we have

$$\mathcal{S}(M_1^c) = \{d+1, \ldots, 9\} \cup \{1, \ldots, d\}^{*\{k+1\}},$$
$$\mathcal{S}(M_2^c) = \{d+1, \ldots, 9\} \cup \{1, \ldots, d\}^{*\{m+1\}}.$$

Thus

$$\frac{\left|\mathcal{S}(M_1^c) \cap \mathcal{S}(M_2^c)\right|}{\left|\mathcal{S}(M_1^c)\right|} = \frac{9-d}{9-d+d^{k+1}} \underset{k\to\infty}{\longrightarrow} 0$$

and this shows the first part of the theorem. If M' contains all digits we have $d = 9$, hence the second part follows. q.e.d.

Example II.1.34 Let $M' = \{1, 2, 33, 44, 55, 66, 77, 88, 99\}$. Then we have $d = 2$ and $\widetilde{M} = \{33, 44, 55, 66, 77, 88, 99\}$. We take $k = 3$ and $m = 5$. Then

$$M_1 = \{1, 2, 11, 12, 21, 22, 111, 112, 122, 121, 211, 212, 221, 222\} \cup X \cup \widetilde{M} \cup Y,$$

where

- X consists of all natural numbers that contain only the digits 1 and 2, and whose number of digits is of the form $5 + 2l, l \in \mathbb{N}_0$,
- Y is the set that contains all natural numbers that contain only the digits $1, 2$, and 0, where the 0 has to occur,

i.e.,

$$\begin{aligned} X = \{&11111, 11112, 11121, 11122, 11211, 11212, 11221, 11222,\\ &12111, 12112, 12121, 12122, 12211, 12212, 12221, 12222,\\ &21111, 21112, 21121, 21122, 21211, 21212, 21221, 21222,\\ &22111, 22112, 22121, 22122, 22211, 22212, 22221, 22222, \ldots\},\\ Y = \{&10, 20, 101, 102, 201, 202, 110, 120, 210, 220, \ldots\}. \end{aligned}$$

We have

$$M_2 = \{1, 2, 11, 12, 21, 22, 111, 112, 122, 121, 211, 212, 221, 222, \ldots\} \cup X \cup \widetilde{M} \cup Y,$$

where the maximal number of digits in the first set is 5. Then

$$\mathcal{S}(M_1) = \{1, 2\} \cup \widetilde{M} = \mathcal{S}(M_2).$$

Further,

$$M_1^c = (\langle\{3,\dots,9\}\rangle \cup A \cup B \cup Z) \setminus \widetilde{M},$$

where

- A consists of all numbers that contain only the digits 1 and 2, and that have exactly 4 digits,
- B consists of all numbers that contain only the digits 1 and 2, and that have exactly 6 digits,
- Z consists of all numbers that contain only the digits 1 and 2, and whose number of digits is of the form $8 + 2l, l \in \mathbb{N}_0$.

Similarly, we have

$$M_2^c = (\langle\{3,\dots,9\}\rangle \cup B \cup Z) \setminus \widetilde{M}.$$

So we get

$$\mathcal{S}(M_1^c) = \{3,\dots,9\} \cup A,\ \ \mathcal{S}(M_2^c) = \{3,\dots,9\} \cup B,\ \ \mathcal{S}(M_1^c) \cap \mathcal{S}(M_2^c) = \{3,\dots,9\},$$

and if we let k and m grow, the sets A and B get bigger.

Note that Theorem II.1.33 is false if M' does not contain any digit, since then $\mathcal{S}(M_1^c) = \mathcal{S}(M_2^c) = \{1,\dots,9\}$.

II.1.6 Heuristics

Since it seems that there are no useful results when considering basic set operations, we could try to get heuristic results. For that, we note that if a minimal set contains "many" natural numbers with a "small" number of digits, then the number of elements in the minimal set is "small", and vice versa. This leads to the following definition

Definition II.1.35 We call a function $\mu : \mathbb{N} \to \mathbb{R}_{>0}$ a **digit measure** if the following properties hold:

- Whenever $\#x < \#y$, we have $\mu(x) > \mu(y)$.
- μ is constant on natural numbers with the same number of digits.

If further $\mu(x) \to 0$ when $\#x \to \infty$, we call μ a **zero digit measure**.

Every digit measure induces a function $\widetilde{\mu} : \mathscr{P}(\mathbb{N}) \to \overline{\mathbb{R}}_{>0} := \mathbb{R}_{>0} \cup \{\infty\}$ by $\widetilde{\mu}(A) := \sum_{a \in A} \mu(a)$. By abuse of notation, we will write $\widetilde{\mu} = \mu$ and also call this a digit measure. We call a digit measure **finite** if $\mu(\mathbb{N}) < \infty$, and **infinite** if $\mu(\mathbb{N}) = \infty$.

Hence a digit measure "measures" a number x by looking only at the number of digits of x.

Example II.1.36 Some seemingly natural digit measures are

$$\mu_1(n) = 10^{-\#n}, \ \mu_2(n) = 10^{-2\#n}, \ \mu_3(n) = \frac{1}{\#n} 10^{-\#n}, \ \mu_4(n) = \frac{1}{(\#n)^2} 10^{-\#n}.$$

We normalize them in the following way:

$$\mu_c := \frac{10}{9}\mu_1, \qquad \mu_g := 10\mu_2, \qquad \mu_h := \frac{10}{9}\mu_3, \qquad \mu_z := \frac{20}{3\pi^2}\mu_4.$$

Then, if we let A_k denote the set of natural numbers with k digits, we have

$$\begin{array}{lll} \mu_c(n) = \frac{10}{9} 10^{-\#n}, & \mu_c(A_k) = 1, & \mu_c(\mathbb{N}) = \infty, \\ \mu_g(n) = 10^{1-2\#n}, & \mu_g(A_k) = 9 \cdot 10^{-k}, & \mu_g(\mathbb{N}) = 1, \\ \mu_h(n) = \frac{10}{9} \frac{1}{\#n} 10^{-\#n}, & \mu_h(A_k) = \frac{1}{k}, & \mu_h(\mathbb{N}) = \infty, \\ \mu_z(n) = \frac{20}{3\pi^2} \frac{1}{(\#n)^2} 10^{-\#n}, & \mu_z(A_k) = \frac{6}{\pi^2} \frac{1}{k^2}, & \mu_z(\mathbb{N}) = 1. \end{array}$$

All these digit measures are zero digit measures.

We will try to use digit measures to obtain results about minimal sets. As we will see what follows, this is nearly as impossible as it was to obtain set-theoretic results.

It would be nice if there was a digit measure μ such that $\frac{\mu(\mathcal{S}(M))}{\mu(M)}$ is constant for all $M \subset \mathbb{N}$. The next theorem says that even a weaker version is not possible.

Theorem II.1.37 *There is no digit measure μ with the following property: There is a function $f : \mathbb{R} \to \mathbb{R}$ such that whenever $\mu(M) = x$ we have $\mu(\mathcal{S}(M)) = f(x)$.*

Proof. We consider the congruence classes modulo 3. Since any such congruence class has the same number of natural numbers with k digits for every k, we have $\mu([a]_3) = \mu([b]_3)$ for every $a, b \in \mathbb{N}$ and for every digit measure. With Theorem II.1.9 (compare also Table A.3.1 in Appendix A.3.2) we get

$$\mu(\mathcal{S}([0]_3)) = 3\mu(1) + 18\mu(10) + 54\mu(100) > 3\mu(1) + 9\mu(10) = \mu(\mathcal{S}([1]_3)).$$

q.e.d.

So even if we knew $\mu(M)$ for a set M (which can be near to impossible to compute exactly for infinite sets), we cannot say anything sharp about $\mu(\mathcal{S}(M))$. We could try to find bounds, i.e., real numbers k_1, k_2, c_1, c_2 such that

$$k_1 \leq \mu(\mathcal{S}(M)) \leq k_2, \qquad c_1 \leq \frac{\mu(\mathcal{S}(M))}{\mu(M)} \leq c_2.$$

If μ is an arbitrary digit measure, then for the first problem we only have the bounds $0 < \mu(\mathcal{S}(M)) < \infty$, since for every $c \in \mathbb{R}_{>0}$ the function $c\mu$ also is a digit measure. So since this question depends on the normalization of μ, we will focus on the second question.

It is clear that $0 \leq \frac{\mu(\mathcal{S}(M))}{\mu(M)} \leq 1$ for every digit measure μ and every set M. We show that there are (in general) no better bounds.

Lemma II.1.38 *Let $M' \subset \mathbb{N}$ be a finite, not truncatable set and μ an arbitrary zero digit measure. Then for any $\varepsilon \in (0,1)$ there is an infinite set M such that $M' = \mathcal{S}(M)$ and*

$$\frac{\mu(\mathcal{S}(M))}{\mu(M)} > 1 - \varepsilon.$$

Proof. Let $m := \mu(M')$ and $R < \frac{m\varepsilon}{1-\varepsilon}$. Let $L \subset \mathbb{N}$ with $L \subset \langle M' \rangle$ and $L \cap M' = \emptyset$. Construct a set $T \subset L$ in the following way: Start with an empty set T and choose $x_1 \in L$ such that $\mu(x_1) < R - \mu(T)$. Add x_1 to T. Now choose $x_2 \in L \backslash \{x_1\}$ such that $\mu(x_2) < R - \mu(T)$ and add x_2 to T. Continuing this process (since μ is a zero digit measure, there is always a suitable x_k) yields an infinite set T with $\mu(T) < R$. Let $M = M' \cup T$. Then we get $\mu(\mathcal{S}(M)) = \mu(M') = m$ and $\mu(M) = \mu(M') + \mu(T) < \frac{m}{1-\varepsilon}$. Hence the lemma follows. q.e.d.

Now we turn our attention to a lower bound. It seems to be more difficult to obtain results about lower bounds than to get those about upper bounds. If μ is an infinite zero digit measure, it is clear that there are no better bounds other than $\frac{\mu(\mathcal{S}(M))}{\mu(M)} \geq 0$. We investigate the problem for the two finite zero digit measures mentioned in Example II.1.36, i.e., μ_g and μ_z. First it is clear that for given M' (finite and not truncatable) and any set M with $\mathcal{S}(M) = M'$ we have

$$\frac{\mu(M')}{\mu(M)} \geq \frac{\mu(M')}{\mu(\langle M' \rangle)}, \tag{II.1.1}$$

so we restrict ourselves to this case. Given any finite not truncatable set M', both sides of the inequality in (II.1.1) are at least $\frac{\mu(M')}{\mu(\mathbb{N})}$. Since $\mu(\mathbb{N})$ is finite, this gives

a lower bound depending on M'. We show that for μ_g and μ_z there are no global lower bounds.

At least for μ_g this is counterintuitive, since for any k we have

$$\frac{\mu_g(\mathcal{S}(\langle A_k\rangle))}{\mu_g(\langle A_k\rangle)} = \frac{\mu_g(A_k)}{\sum_{i\geq k}\mu_g(A_i)} = \frac{9\cdot 10^{-k}}{\sum_{i\geq k} 9\cdot 10^{-i}} = \frac{9}{10}.$$

For μ_z this ratio becomes more complicated and depends on k:

$$\frac{\mu_z(\mathcal{S}(\langle A_k\rangle))}{\mu_z(\langle A_k\rangle)} = \frac{\frac{6}{\pi^2}\frac{1}{k^2}}{\sum_{i\geq k}\frac{6}{\pi^2}\frac{1}{i^2}} = \frac{\frac{1}{k^2}}{\sum_{i\geq k}\frac{1}{i^2}} = \frac{\frac{1}{k^2}}{\frac{\pi^2}{6} - \sum_{i=1}^{k-1}\frac{1}{i^2}}.$$

Figure II.1.1 shows the ratio $\frac{\mu_z(\mathcal{S}(\langle A_k\rangle))}{\mu_z(\langle A_k\rangle)}$ for $k \in \{1,\ldots,10\}$.

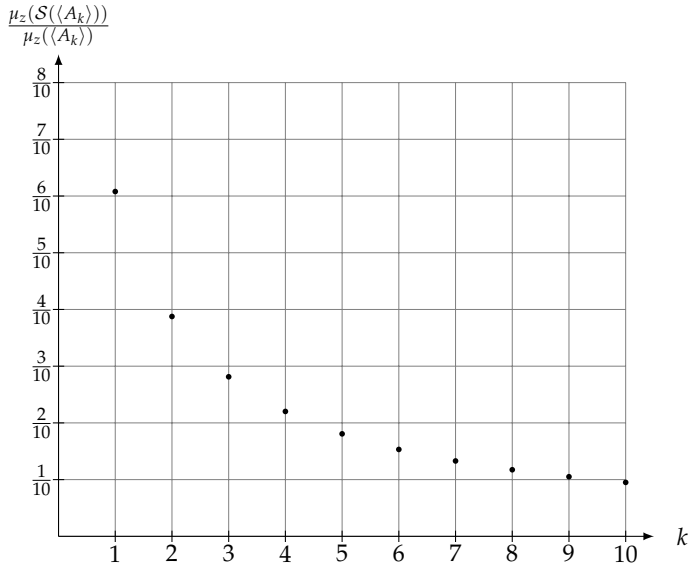

Figure II.1.1: Values for $\frac{\mu_z(\mathcal{S}(\langle A_k\rangle))}{\mu_z(\langle A_k\rangle)}$.

Due to the values in Figure II.1.1 we can already guess that the ratio tends to zero.

Theorem II.1.39 *We have*

$$\frac{\mu_z(A_k)}{\mu_z(\langle A_k\rangle)} \underset{k\to\infty}{\longrightarrow} 0.$$

Proof. This follows from

$$\frac{\mu_z(A_k)}{\mu_z(\langle A_k\rangle)} = \frac{\frac{1}{k^2}}{\sum_{i\geq k}\frac{1}{i^2}} < \frac{\frac{1}{k^2}}{\sum_{i=k}^{2k}\frac{1}{i^2}} < \frac{\frac{1}{k^2}}{\sum_{i=k}^{2k}\frac{1}{(2k)^2}} = \frac{\frac{1}{k^2}}{\frac{1}{4k^2}(k+1)} = \frac{4}{k+1}.$$

q.e.d.

Now we consider μ_g. Here we cannot take the sets A_k, since the ratio would be constant as seen above.

Theorem II.1.40 *Let S_k denote any set with $S_k = \{n\}$ such that $\#n = k$. Then*

$$\frac{\mu_g(S_k)}{\mu_g(\langle S_k \rangle)} \underset{k\to\infty}{\longrightarrow} 0.$$

Proof. First we need to examine $\mu(\langle S_k \rangle)$. Unfortunately, we cannot determine this value exactly, but we can give a bound which suffices for our purpose. Our aim is to construct a subset of $\langle S_k \rangle$ by constructing elements that have exactly $l+1$ digits from those that have l digits. Choose an element $x \in \langle S_k \rangle$ with l digits. Then we have $l+1$ positions where we can insert a new digit to form a number with $l+1$ digits. To make sure that the constructed numbers are all distinct, we choose the new digit such that its two neighbours are different from the new digit. Thus we have for any position at least 7 possibilites to do so. Note that we only use one l-digit number in this process, since otherwise we could get the same $l+1$-digit number more than once. This construction gives a subset T of $\langle S_k \rangle$, where we are missing some (in fact a lot of) elements: For $i \in \mathbb{N}$, T has exactly $7(k+i)$ elements that have exactly $k+i$ digits. We get

$$\begin{aligned}
\frac{\mu_g(S_k)}{\mu_g(\langle S_k \rangle)} &< \frac{\mu_g(S_k)}{\mu_g(T)} \\
&= \frac{10^{1-2k}}{10^{1-2k} + \sum_{i=1}^{\infty} 7(k+i)10^{1-2(k+i)}} \\
&= \frac{10^{1-2k}}{10^{1-2k} + 7k \cdot 10^{1-2k} \sum_{i=1}^{\infty} 100^{-i} + 7 \cdot 10^{1-2k} \cdot \sum_{i=1}^{\infty} i \cdot 100^{-i}} \\
&= \frac{10^{1-2k}}{10^{1-2k} + 7 \cdot 10^{1-2k} \cdot \left(\frac{k}{99} + \frac{100}{9801}\right)} \\
&\underset{k\to\infty}{\to} 0.
\end{aligned}$$

q.e.d.

In fact the above proof works for a zero digit measure μ if and only if

$$\sum_{i=1}^{\infty} (k+i) \frac{\mu^{k+i}}{\mu^k} \underset{k\to\infty}{\longrightarrow} \infty, \tag{II.1.2}$$

where μ^l denotes the value of μ at any natural number with l digits.

Example II.1.41 Consider the zero digit measure $\mu(n) = 10^{-(\#n)^2}$. Then

$$\begin{aligned}\sum_{i=1}^{\infty}(k+i)\frac{\mu^{k+i}}{\mu^k} &= \sum_{i=1}^{\infty}(k+i)10^{-i^2-2ki} < k\sum_{i=1}^{\infty}\left(10^{-2k}\right)^i + \sum_{i=1}^{\infty} i\left(10^{-2k}\right)^i \\ &= k\frac{10^{-2k}}{1-10^{-2k}} + \frac{10^{-2k}}{(1-10^{-2k})^2} \\ &= \frac{k\cdot 10^{-2k}\cdot(1-10^{-2k}) + 10^{-2k}}{(1-10^{-2k})^2} \\ &\to 0,\end{aligned}$$

hence μ does not satisfy the condition in (II.1.2). When considering the sets A_k, we get $\mu(A_k) = 9\cdot 10^{-k^2+k-1}$ and

$$\begin{aligned}\mu(\langle A_k\rangle) &= \sum_{i=0}^{\infty} 9\cdot 10^{k+i-1}\cdot 10^{-(k+i)^2} = 9\cdot 10^{-k^2+k-1}\sum_{i=0}^{\infty} 10^{-i^2+i-2ki} \\ &= \mu(A_k)\left(1+\underbrace{\sum_{i=1}^{\infty} 10^{-i^2+i-2ki}}_{\to 0}\right),\end{aligned}$$

thus $\frac{\mu(A_k)}{\mu(\langle A_k\rangle)} \underset{k\to\infty}{\longrightarrow} 1$. It would be interesting to know whether there are lower bounds for this digit measure.

We conclude this section with the remark that there cannot be any results that relate the minimal set of M to an asymptotic formula or the density of M. The reason is that both an asymptotic formula and the densitiy of M would not change if we changed a finite number of elements in M, whereas the minimal set could change drastically through this.

II.1.7 An Odd End

Choosing a base other than 10 would not make things easier in general (see [BDS16] for recent results in this direction). Of course, with the binary expansion in place of the decimal expansion, a few minimal sets would look more simple, for example $\mathcal{S}(\mathbb{P}) = \{\langle 10\rangle_2, \langle 11\rangle_2\}$. However, there are still plenty of sets, interesting from a number-theoretical point of view, where the corresponding minimal set seems to be difficult to describe.

For instance, the set of perfect numbers is conjectured to contain no odd numbers, and every even perfect number can be written in the form $2^{p-1}(2^p-1)$

(compare Theorem 0.5.1). Lucas [Luc90] showed that every even perfect number different from 6 and 496 ends with decimal digits $16, 28, 36, 56$, or 76. Hence, if there is no odd perfect number, then the minimal set of the set of perfect numbers is given by $\{6, 28\}$. If there exists an odd perfect number, although much can be said about its hypothetical multiplicative structure, we cannot exclude the possibility that this odd number is a minimal element. In base 2, the minimal set would be $\{\langle 110\rangle_2\}$ if there was no odd perfect number, and we leave the computation of the minimal set with respect to other bases (under the same assumption) to the interested reader.

II.1.8 Future Work

Since this topic is not well known, there is some more possible work. Of course one could try to determine the minimal sets for more (arithmetically interesting) sets, in different bases.

- Determine the minimal set (in different bases) for further sets.

In particular it would be nice to see if Shallit's conjecture about the minimal set of the powers of 2 is correct.

- Determine the minimal set of $\{2^n : n \in \mathbb{N}_0\}$.

In this chapter we formulated three more conjectures which one can give some attention.

- Prove or disprove Conjectures II.1.3, II.1.19, and II.1.24.

Concerning bounds for $\frac{\mu(\mathcal{S}(M))}{\mu(M)}$ we have only considered zero digits measures. One could examine this problem for other digit measures (for example the one in Example II.1.41):

- Let μ be a digit measure. If μ is not a zero digit measure, are there bounds for $\frac{\mu(\mathcal{S}(M))}{\mu(M)}$?
- Find a lower bound for the digit measure in Example II.1.41.

Apart from these conjectures and the determination of some more minimal sets, it would be great to know if Algorithm II.1.2 can be applied only to unions of congruence classes or if this algorithm can give the minimal sets for other sets, too.

- Characterize the truncating stable partitions A_i of $\mathbb{N}$ (compare Definition II.1.15).

II.2

Adding Generators in Abelian Groups

In this chapter we examine a connection between graph theory and number theory that is given by Cayley graphs. More precisely, we will determine the neighbourhood of the neighbourhood of vertices in some Cayley graphs of abelian groups. This can be done by examining the sumset of atoms in abelian groups.

The results of this chapter have been published in [Kre15a].

II.2.1 Introduction

In [SS13], Sander and Sander considered the problem of finding the neighbourhood of the neighbourhood of the vertex 0 in gcd graphs $X_n(D)$, i.e., in Cayley graphs $X(G, S)$ for cyclic groups G and certain subsets $S \subset G$ (compare Chapter I.5). This problem reduces to the problem of adding generators in cyclic groups. In this chapter, we generalize their result to the case of abelian groups and take a short look at non-abelian groups. Throughout this chapter, all groups are finite and additively written.

In their work, Sander and Sander treat the set

$$S_{n;a,b}(c) = \{(u, v) \in \mathrm{atom}(a) \times \mathrm{atom}(b) : u + v = c\}$$

for given $a, b, c \in \mathbb{Z}/n\mathbb{Z}$, where

$$\begin{aligned}\mathrm{atom}(a) &:= \{a' \in \mathbb{Z}/n\mathbb{Z} : \langle a' \rangle = \langle a \rangle\} = a(\mathbb{Z}/n\mathbb{Z})^* \\ &= \{au : 1 \leq u \leq n, \gcd(u, n) = 1\}.\end{aligned}$$

They derive a representation of $\mathrm{atom}(a)$, namely

$$\mathrm{atom}(a) = \{ax : 1 \leq x \leq \mathrm{ord}(a), \gcd(x, \mathrm{ord}(a)) = 1\} =: (a)_n^\star,$$

where each element $b \in \mathrm{atom}(a)$ is uniquely represented in $(a)_n^\star$, i.e., for each $b \in \mathrm{atom}(a)$ there is exactly one x with $1 \leq x \leq \mathrm{ord}(a), \gcd(x, \mathrm{ord}(a)) = 1$ such that $b = ax$. Here (and in the rest of this chapter) $\mathrm{ord}(a)$ denotes the order of a in $(\mathbb{Z}/m\mathbb{Z}, +)$ for a given $m \geq 2$, i.e., $\mathrm{ord}(a) = \frac{a}{\gcd(a,m)}$.

It has been known before (see [KS12]), that the sum

$$\mathrm{atom}(a) + \mathrm{atom}(b) = \{a' + b' : a' \in \mathrm{atom}(a), b' \in \mathrm{atom}(b)\}$$

is the disjoint union of atoms, even in the case of abelian groups. In their paper, Sander and Sander determine which atoms will appear in the union and they establish a formula for $N_{n;a,b}(c) := |S_{n;a,b}(c)|$. More precisely, they show (among others):

Theorem II.2.1 *([SS13]) Let $n \in \mathbb{N}$ with divisors a and b, let $g := \gcd(a, b)$, and let $c \in \mathbb{Z}/n\mathbb{Z}$. We set $n' := \frac{n}{g}, a' := \frac{a}{g}$, and $b' := \frac{b}{g}$. Let $m = \frac{n'}{a'b'}$ and $\widetilde{m}_3$ be the largest divisor of m with $\gcd(\widetilde{m}_3, a'b') = 1$. If $g|c$, let further $c' = \frac{c}{g}$.*

1. *If $2 \nmid n'$ or $2|a'b'$, then*

   (a) *$N_{n;a,b}(c) > 0$ if and only if $g|c$ and $\gcd(c', a'b') = 1$.*

   (b) *We have*
   $$(a)_n^\star + (b)_n^\star = \bigcup_{d|\widetilde{m}_3} g(d)_n^\star.$$

2. *If $2|n'$ and $2 \nmid a'b'$, then*

   (a) *$N_{n;a,b}(c) > 0$ if and only if $g|c$, c' is even, and $\gcd(c', a'b') = 1$.*

   (b) *We have*
   $$(a)_n^\star + (b)_n^\star = \bigcup_{d|\widetilde{m}_3, 2|d} g(d)_n^\star.$$

In both cases, we have

$$N_{n;a,b}(c) = m \prod_{p|m,p|a'b'} \left(1 - \frac{1}{p}\right) \prod_{p|n',p\nmid a'b',p|c'} \left(1 - \frac{1}{p}\right) \prod_{p|n',p\nmid a'b'c'} \left(1 - \frac{2}{p}\right)$$

if $g|c$ *and* $\gcd(c', a'b') = 1$.

There has been some other work in this area of research. Alperin and Peterson [AP12] studied atomic numbers, i.e., the number of atoms in some boolean algebra for some groups. Sun and Yang [SY14] generalized the result of Sander and Sander by examining sums of t atoms for arbitrary $t \geq 2$, the case $t = 2$ being the case covered in [SS13].

In this chapter, we will take a look at how the result of Sander and Sander generalizes to abelian groups. We will show that the decomposition of abelian groups in cyclic groups yields an immediate generalization of Theorem II.2.1 to abelian groups.

II.2.2 Abelian groups

Let us first generalize the definitions and the problem to the new context. All the definitions and notations here are completely analogous to those in [SS13].

Let A be an abelian group. We know that A can be decomposed into cyclic groups, say

$$A \cong \bigtimes_{i=1}^{k} \mathbb{Z}/m_i\mathbb{Z}.$$

There are ways to restrict the orders m_i of the cyclic groups to make this decomposition unique, but since we do not need this we will not pose any conditions on the m_i. We will use this decomposition to regard A not only as an abelian group but also as a ring.

As for cyclic groups, the **atom** of $a \in A$ is the set of generators of the subgroup $\langle a \rangle$, i.e.,

$$\operatorname{atom}(a) := \{a' \in A : \langle a' \rangle = \langle a \rangle\} = aA^*.$$

For an $a \in A$, let $a = (a_1, \ldots, a_k)$ be the corresponding element in $\bigtimes_{i=1}^{k} \mathbb{Z}/m_i\mathbb{Z}$. Then we get

$$\operatorname{atom}(a) \cong \{(a_1u_1, \ldots, a_ku_k) : 1 \leq u_i \leq m_i, \gcd(u_i, m_i) = 1 \text{ for all } i\}.$$

This is correct since

$$\begin{aligned} A^* &\cong \bigtimes_{i=1}^{k} \{u_i \in \mathbb{Z}/m_i\mathbb{Z} : \gcd(u_i, m_i) = 1\} \\ &= \{(u_1, \ldots, u_k) : 1 \le u_i \le m_i, \gcd(u_i, m_i) = 1 \text{ for all } i\}. \end{aligned}$$

As in the case of cyclic groups, there are elements in $\mathrm{atom}(a)$ which are represented more than once, so we wish to find a set that represents any of its elements exactly once. Let

$$\begin{aligned} (a)^\star_A &:= \{(a_1x_1, \ldots, a_kx_k) : 1 \le x_i \le \mathrm{ord}(a_i), \gcd(x_i, \mathrm{ord}(a_i)) = 1 \text{ for all } i\} \\ &= \{(v_1, \ldots, v_k) : v_i \in (a_i)^\star_{m_i}\}. \end{aligned}$$

Here $\mathrm{ord}(a_i)$ denotes the order of a_i in $(\mathbb{Z}/m_i\mathbb{Z}, +)$, i.e., $\mathrm{ord}(a_i) = \frac{m_i}{\gcd(a_i, m_i)}$. Then we have

$$\begin{aligned} \mathrm{atom}(a) &\cong \{(a_1u_1, \ldots, a_kx_k) : 1 \le u_i \le m_i, \gcd(u_i, m_i) = 1 \text{ for all } i\} \\ &= \bigtimes_{i=1}^{k} \{a_iu_i : 1 \le u_i \le m_i, \gcd(u_i, m_i) = 1\} \\ &= \bigtimes_{i=1}^{k} \{a_ix_i : 1 \le x_i \le \mathrm{ord}(a_i), \gcd(x_i, \mathrm{ord}(a_i)) = 1\} \\ &= \{(a_1x_1, \ldots, a_kx_k) : 1 \le x_i \le \mathrm{ord}(a_i), \gcd(x_i, \mathrm{ord}(a_i)) = 1 \text{ for all } i\} \\ &= (a)^\star_A \end{aligned}$$

and by definition, we get

$$(a)^\star_A + (b)^\star_A = \{(v_1, \ldots, v_k) : v_i \in (a_i)^\star_{m_i} + (b_i)^\star_{m_i}\}.$$

Together with Theorem II.2.1 this gives the sumset $(a)^\star_A + (b)^\star_A$ for abelian groups, see also Section II.2.3.

For $c = (c_1, \ldots, c_k) \in A$ we define

$$\begin{aligned} S_{A;a,b}(c) &= \{(u, v) \in (a)^\star_A \times (b)^\star_A : u + v = c\} \\ &= \{(u, v) \in (a)^\star_A \times (b)^\star_A : u_i + v_i = c_i \text{ for all } i\} \end{aligned}$$

and $N_{A;a,b}(c) := |S_{A;a,b}(c)|$. Then we have the following result.

Theorem II.2.2 *Let A be an abelian group with $A \cong \bigtimes_{i=1}^{k} \mathbb{Z}/m_i\mathbb{Z}$ and $a, b, c \in A$. For $x \in A$ let $(x_1, \ldots, x_k)$ be the corresponding element in $\bigtimes_{i=1}^{k} \mathbb{Z}/m_i\mathbb{Z}$. Then*

$$S_{A;a,b}(c) \cong \bigtimes_{i=1}^{k} S_{m_i;a_i,b_i}(c_i) \quad \text{and} \quad N_{A;a,b}(c) = \prod_{i=1}^{k} N_{m_i;a_i,b_i}(c_i).$$

Proof. This follows immediately from the corresponding result for cyclic groups and the decomposition of A into cyclic groups:

$$\begin{aligned}
S_{A;a,b}(c) = \{((a_1x_1, \ldots, a_kx_k), (b_1y_1, \ldots, b_ky_k)) : \\
&1 \le x_i \le \operatorname{ord}(a_i), \gcd(x_i, \operatorname{ord}(a_i)) = 1, \\
&1 \le y_i \le \operatorname{ord}(b_i), \gcd(y_i, \operatorname{ord}(b_i)) = 1, \\
&a_ix_i + b_iy_i \equiv c_i \bmod m_i \text{ for all } i\} \\
\cong \{(a_1x_1, b_1y_1) : \\
&1 \le x_1 \le \operatorname{ord}(a_1), \gcd(x_1, \operatorname{ord}(a_1)) = 1, \\
&1 \le y_1 \le \operatorname{ord}(b_1), \gcd(y_1, \operatorname{ord}(b_1)) = 1, \\
&a_1x_1 + b_1y_1 \equiv c_1 \bmod m_1\} \\
\times \cdots \times \{(a_kx_k, b_ky_k) : \\
&1 \le x_k \le \operatorname{ord}(a_k), \gcd(x_k, \operatorname{ord}(a_k)) = 1, \\
&1 \le y_k \le \operatorname{ord}(b_k), \gcd(y_k, \operatorname{ord}(b_k)) = 1, \\
&a_kx_k + b_ky_k \equiv c_k \bmod m_k\} \\
= \bigtimes_{i=1}^{k} S_{m_i;a_i,b_i}(c_i).
\end{aligned}$$

q.e.d.

II.2.3 Applications for Cayley graphs

As already mentioned, we can use the results about sumsets to deduce properties of neighbourhoods in Cayley graphs. Recall that the Cayley graph $X(H, S)$ for a finite (additive) group H and a subset $S \subset H$ with $0 \notin S, -S = S$ is defined as follows:

The vertices of $X(H, S)$ are the elements of H, and two elements $g, h \in H$ are adjacent whenever $g - h \in S$.

We are interested in the sets $N_1(v)$ and $N_2(v)$, i.e., in those vertices with distance 1 or 2 from v. Since Cayley graphs are transitive (compare Paragraph 0.4.5), it suffices to consider $v = 0$. In this case we immediately get $N_1(v) = S$ and $N_2(v) = (S+S)\backslash\{0\}$ if $X(H,S)$ has no triangles, so our result explicitly yields the set $N_2(v)$. In our example, we will examine a set S which is the disjoint union of atoms rather than being an atom itself.

Example II.2.3 Let $H = \mathbb{Z}/30\mathbb{Z} \times \mathbb{Z}/12\mathbb{Z}$ and $S = \mathrm{atom}(3,2) \cup \mathrm{atom}(5,3)$. Since

$$(3)^\star_{30} = \{3,9,21,27\}, \quad (5)^\star_{30} = \{5,25\}, \quad (2)^\star_{12} = \{2,10\}, \quad (3)^\star_{12} = \{3,9\}$$

we have

$$\begin{aligned} N_0(0,0) =& \{(0,0)\}, \\ N_1(0,0) =& \{(3,2),(3,10),(9,2),(9,10),(21,2),(21,10),(27,2),(27,10), \\ & (5,3),(5,9),(25,3),(25,9)\}. \end{aligned}$$

To determine $N_2(0,0)$ we compute the sums of the atoms. We get

$$\begin{aligned} (3)^\star_{30} + (3)^\star_{30} &= \bigcup_{2|d,d|10} 3(d)^\star_{30} = (6)^\star_{30} \cup (0)^\star_{30}, \\ (3)^\star_{30} + (5)^\star_{30} &= \bigcup_{2|d,d|2} (d)^\star_{30} = (2)^\star_{30}, \\ (5)^\star_{30} + (5)^\star_{30} &= \bigcup_{2|d,d|6} 5(d)^\star_{30} = (10)^\star_{30} \cup (0)^\star_{30}, \\ (2)^\star_{12} + (2)^\star_{12} &= \bigcup_{2|d,d|6} 2(d)^\star_{12} = (4)^\star_{12} \cup (0)^\star_{12}, \\ (2)^\star_{12} + (3)^\star_{12} &= \bigcup_{d|1} (d)^\star_{12} = (1)^\star_{12}, \\ (3)^\star_{12} + (3)^\star_{12} &= \bigcup_{2|d,d|4} 3(d)^\star_{12} = (6)^\star_{12} \cup (0)^\star_{12}, \end{aligned}$$

and

$$\begin{aligned} &(2)^\star_{30} = \{2,4,8,14,16,22,26,28\}, \quad (6)^\star_{30} = \{6,12,18,24\}, \quad (10)^\star_{30} = \{10,20\}, \\ &(1)^\star_{12} = \{1,5,7,11\}, \qquad (4)^\star_{12} = \{4,8\}, \qquad (6)^\star_{12} = \{6\}. \end{aligned}$$

Thus the possible vertices with distance 2 are

$$\begin{aligned} Y = \{&(2,1),(4,1),(8,1),(14,1),(16,1),(22,1),(26,1),(28,1),\\ &(2,5),(4,5),(8,5),(14,5),(16,5),(22,5),(26,5),(28,5),\\ &(2,7),(4,7),(8,7),(14,7),(16,7),(22,7),(26,7),(28,7),\\ &(2,11),(4,11),(8,11),(14,11),(16,11),(22,11),(26,11),(28,11),\\ &(6,4),(12,4),(18,4),(24,4),(6,8),(12,8),(18,8),(24,8),\\ &(6,0),(12,0),(18,0),(24,0),(0,4),(0,8),\\ &(10,6),(20,6),(10,0),(20,0),(0,6)\}. \end{aligned}$$

Since none of them is a neighbour of $(0,0)$, we indeed have $N_2(0,0) = Y$.

II.2.4 Non-abelian groups

As we have seen, the sum of a disjoint union of atoms is again a disjoint union of atoms when H is abelian. If H is non-abelian, this is not true in general, as the following example shows.

Example II.2.4 Let $H = \mathrm{GL}_2(\mathbb{F}_7)$ and $S = \mathrm{atom}(X) \cup \mathrm{atom}(Y)$ where

$$X = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, Y = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}.$$

Then X has order 4 and Y has order 6, hence we have $\mathrm{atom}(X) = \{X, X^3\}$ and $\mathrm{atom}(Y) = \{Y, Y^5\}$. So we get

$$S = \mathrm{atom}(X) \cup \mathrm{atom}(Y) = \left\{ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} \right\},$$

$$S \cdot S = \{X^2, Y^2, Y^4, YX, Y^5X, YX^3, Y^5X^3, XY, XY^5, X^3Y, X^3Y^5\}$$

$$= \left\{ \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 3 & -3 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 3 \\ -2 & 3 \end{pmatrix}, \begin{pmatrix} -2 & 1 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} -2 & -1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 2 & -1 \\ 1 & -1 \end{pmatrix}, \right.$$

$$\left. \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 1 & -2 \end{pmatrix}, \begin{pmatrix} -1 & -1 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ -1 & 2 \end{pmatrix} \right\}.$$

Let $A := \begin{pmatrix} -2 & 1 \\ -1 & 1 \end{pmatrix}$. A has order 16, hence

$$\text{atom}(A) = \{A, A^3, A^5, A^7, A^9, A^{11}, A^{13}, A^{15}\}.$$

But $A^3 = \begin{pmatrix} 2 & 2 \\ -2 & 1 \end{pmatrix}$, therefore we have $A \in S \cdot S$ and $A^3 \notin S \cdot S$, thus $S \cdot S$ cannot be the union of atoms.

In the context of Cayley graphs this means that the set $N_2(0)$ need not be a disjoint union of atoms if S is a disjoint union of atoms.

II.2.5 Future Work

There are a few open questions which one could consider. When searching for an example in the non-abelian case, one finds some situations in which the set $S \circ S$ is a disjoint union of atoms if S is a disjoint union of atoms, for example when examining symmetric groups. So one could ask:

- Is $S \circ S$ a disjoint union of atoms if S is a disjoint union of atoms and S is a subset of some symmetric group?

In Example II.2.4, the set S was the disjoint union of two atoms. What happens if we take S to be an atom itself?

- Is $S \circ S$ a disjoint union of atoms if S is an atom and $S \subset H$ for a non-abelian group H?

If the above questions will be answered in the negative, this could raise the question of how a set S has to look like such that $S \circ S$ is a disjoint union of atoms.

- Given a non-abelian group H, characterize the subsets $S \subset H$ such that $S \circ S$ is a disjoint union of atoms whenever S is a disjoint union of atoms.

Regarding generalizations of our result, one could take a look at the paper of Sun and Yang [SY14] who examined sums of t atoms in cyclic groups. This should generalize to abelian groups.

- Does an analogous statement of Theorem II.2.2 hold if we consider sums of $t \geq 2$ atoms in abelian groups?

II.3

On the Number of Solutions of Linear Equations over Factor Rings of Principal Ideal Domains

In this chapter we examine certain equations over principal ideal domains R. In the case $R = \mathbb{Z}$ this can be applied in the circle method. We will determine the number of solutions for two different equations, both defined over principal ideal domains R with some finiteness conditions. We will use algebraic methods, thus this chapter shows a connection between algebra and number theory. Some examples for the results of this chapter (in particular of Section II.3.5) can be found in Appendix A.4.

The results of this chapter have been published in [Kre16].

II.3.1 Introduction

We determine the number of solutions of two different kinds of linear equations, namely

$$\{\mathbf{x} \in (R/\mathfrak{a})^n : A\mathbf{x} \in \mathfrak{a}\} \qquad \text{and} \qquad \{\mathbf{x} \in (R/\mathfrak{a})^n : \langle \mathbf{d}, \mathbf{x} \rangle \in \mathfrak{a}\}$$

where R is a principal ideal domain, $\mathfrak{a} \lhd R$ is an ideal of R, $A \in \mathcal{M}_{m,n}(R)$ and $\mathbf{d} \in R^n$ for some $m, n \in \mathbb{N}$. Here $\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n$ if $\mathbf{x} = (x_1, \ldots, x_n) \in R^n$, $\mathbf{y} = (y_1, \ldots, y_n) \in R^n$.

To avoid confusion, throughout this chapter let p denote a prime in $\mathbb{N}$ and π a prime in R (since we are working over principal ideal domains this is equivalent to saying that π is irreducible). If $a \in \mathbb{Z}$, $\prod_{p|a}$ denotes, as usual, the product over all primes (in $\mathbb{N}$) dividing a, while if $a \in R$, the notation $\prod_{\pi|a}$ means the following: Fix a prime decomposition of a in R. Then the prime π contributes to the product if and only if it appears in this fixed decomposition. Whenever $\mathbf{x} \in R^m$ for some m, the condition $\mathbf{x} \in \mathfrak{a}$ for an ideal $\mathfrak{a} \lhd R$ means that each of the components of $\mathbf{x}$ lies in $\mathfrak{a}$. Analogously, $\mathbf{x} \equiv \mathbf{y} \bmod a$ for $\mathbf{x}, \mathbf{y} \in R^m$ means that $x_i \equiv y_i \bmod a$ for the components of $\mathbf{x}$ and $\mathbf{y}$. We write $a \sim b$ for $a, b \in R$ if there is a $c \in R^*$ such that $a = cb$. If $R = \mathbb{Z}$ and the matrix A is a square matrix, we let $\Delta = |\det(A)|$.

Since R is a principal ideal domain, we determine the following numbers:

$$|\{\mathbf{x} \in (R/(a))^n : A\mathbf{x} \in (a)\}| \tag{II.3.1}$$

and

$$|\{\mathbf{x} \in (R/(a))^n : \langle \mathbf{d}, \mathbf{x} \rangle \in (a)\}| \tag{II.3.2}$$

for an $a \in R$.

In the case $R = \mathbb{Z}$, the equations appearing in (II.3.1) and (II.3.2) are just linear congruences, i.e., we determine

$$|\{A\mathbf{x} \equiv \mathbf{0} \bmod a\}| \tag{II.3.3}$$

and

$$|\{\langle \mathbf{d}, \mathbf{x} \rangle \equiv 0 \bmod a\}|\,. \tag{II.3.4}$$

The equation in (II.3.2) is in fact just a special case of the one in (II.3.1). We consider this as a problem in its own right since we will use a more direct approach to solve this (compare also the comments in Section II.3.6). In fact, we shall impose another condition on the vector $\mathbf{d}$ in (II.3.2) and (II.3.4) that will be motivated from the solution to the first problem.

To avoid trivial cases we will always assume that $\mathfrak{a}$ is a nontrivial ideal, i.e., $a \neq 0$ and $a \notin R^*$.

Equations involving matrices have been widely studied. The equation $AX = B$ for given matrices A, B has (among others) been studied by Khatri and Mitra [KM76], Horn, Sergeichuk, and Shaked-Monderer [HSSM05], Don [Don87] and Hua [Dai90]. In their works, the main goal was to find a solution with given properties (for example hermitian, nonnegative definite). Other equations that have been studied include the equations $AXC = B$ (Zhang [Zha04], Khatri and Mitra [KM76], Hua [Dai90]) or systems of matrix equations (Tian [Tia02], Khatri and Mitra [KM76]).

Camon, Levy, and Mann [CLM71] considered the equation $A\mathbf{x} = \mathbf{1}$ over a commutative ring R. In this chapter we study a similar equation, the main goal beeing not to give a criterion for solvability (since our equation will always be solvable), but to find the number of solutions.

The result for the number (II.3.3) (in particular an upper bound) is useful for applications in number theory.

To be more precise, let F be a nonsingular (i.e., the matrix associated to F has nonzero determinant) quadratic form in n variables. To estimate representation numbers of F, one can use a form of the circle method (compare Chapter I.10) and consider the sum (compare [HB96])

$$S_q(\mathbf{c}) = \sum_{a \bmod q}^{*} \sum_{\mathbf{b} \bmod q} e_q(aF(\mathbf{b}) + \langle \mathbf{c}, \mathbf{b} \rangle).$$

Here $q \in \mathbb{N}, \mathbf{c} \in \mathbb{Z}^n$ and $e_q(x)$ is defined as $e_q(x) = \exp(2\pi i x/q)$. The star in the first sum indicates that we only sum over those a coprime to q. In the second sum, every entry of $\mathbf{b}$ runs modulo q independently. An easy estimate using the Cauchy-Schwarz inequality yields

$$\left|S_q(\mathbf{c})\right|^2 \leq \varphi(q) \sum_{a \bmod q}^{*} \sum_{\mathbf{w} \bmod q} e_q(aF(\mathbf{w}) + \langle \mathbf{c}, \mathbf{w} \rangle) \sum_{\mathbf{v} \bmod q} e_q(\langle a\mathbf{v}, \nabla F(\mathbf{w}) \rangle),$$

where φ denotes the Euler totient function and ∇F the gradient of the function F.

Here the last summation will only contribute if $q | \nabla F(\mathbf{w})$. But $\nabla F(\mathbf{w}) = 2M\mathbf{w}$ where M is the underlying matrix of the quadratic form F. To give an explicit bound for $S_q(\mathbf{c})$ we therefore need to determine the number of $\mathbf{w} \in (\mathbb{Z}/q\mathbb{Z})^n$ such that $2M\mathbf{w} \equiv 0 \bmod q$.

II.3.2 The number $|\{\mathbf{x} \in (R/(a))^n : A\mathbf{x} \in (a)\}|$

Since we are working over a principal ideal domain, we can compute the Smith normal form of A, cf. Theorem 0.4.8. Let B denote the Smith normal form of A. Then there exist $U \in \mathrm{GL}_m(R), V \in \mathrm{GL}_n(R)$ such that $A = UBV$. Since U and V are invertible, we see at once that

$$|\{\mathbf{x} \in (R/(a))^n : A\mathbf{x} \in (a)\}| = |\{\mathbf{x} \in (R/(a))^n : B\mathbf{x} \in (a)\}|.$$

So it remains to examine the problem for matrices B in Smith normal form, i.e., we have

$$B = \begin{pmatrix} d_1 & & & & & & \\ & d_2 & & & & & \\ & & \ddots & & & & \\ & & & d_z & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix} \text{ with } d_i | d_{i+1}, i = 1, \ldots, z-1.$$

We will write this in short as $B = \mathrm{diag}(d_1, \ldots, d_r)$ with $r = \min(m, n)$ by allowing $0|0$. The elements d_i are determined up to units $c \in R^*$. Note that $\det(U) \in R^*$ and $\det(V) \in R^*$, thus for $m = n$ we have

$$\det(A) \sim \det(B) = \prod_{i=1}^{n} d_i.$$

We have the following theorem:

Theorem II.3.1 *Let R be a principal ideal domain such that $|R/(\pi)|$ is finite for each irreducible $\pi \in R$. Let $a \in R$ and denote by $v_\pi(h)$ the π-adic valuation of $h \in R$, i.e., the biggest exponent of π that occurs in a prime decomposition of h. Let further $A \in \mathcal{M}_{m,n}(R)$ with Smith normal form $B = \mathrm{diag}(d_1, \ldots, d_r)$ where $r = \min(m, n)$. Then*

$$|\{\boldsymbol{x} \in (R/(a))^n : A\boldsymbol{x} \in (a)\}| = \prod_{\pi | a} \prod_{i=1}^{r} |R/(\pi)|^{\min(v_\pi(d_i), v_\pi(a))}.$$

Proof. Let

$$a \sim \prod_{i=1}^{k} \pi_i^{t_i}$$

be a decomposition of a into primes in R. Then the Chinese remainder theorem yields

$$R/(a) \cong \bigtimes_{i=1}^{k} R/(\pi_i^{t_i}),$$

hence

$$B\mathbf{x} \in (a) \Leftrightarrow B\mathbf{x} \in (\pi_i^{t_i}) \text{ for all } i.$$

Since B is in Smith normal form, this condition means that $d_l x_l \in (\pi^t)$ for all $l = 1, \ldots, m$ for each π (here we omit the index i to avoid too much nested indices and exponents).

We consider each of the equations $d_l x_l \in (\pi^t)$ separately. Write $d_l = \pi^{v_\pi(d_l)} r_l$ with $r_l \in R, \pi \nmid r_l$, i.e., $d_l x_l \in (\pi^{v_\pi(d_l)})$. We distinguish two cases:

- Suppose that $v_\pi(d_l) \geq t$. In this case, $(\pi^{v_\pi(d_l)}) \subset (\pi^t)$, so we can choose x_l arbitrarily in $R/(\pi^t)$.

- Suppose that $v_\pi(d_l) \leq t$. Cancelling $\pi^{v_\pi(d_l)}$ in the equation $d_l x_l \in (\pi^{v_\pi(d_l)})$ yields $r_l x_l \in (\pi^{t-v_\pi(d_l)})$. Since $\pi \nmid r_l$, this means we can choose x_l arbitrarily in $\pi^{t-v_\pi(d_l)} R/(\pi^t) \cong R/(\pi^{v_\pi(d_l)})$.

The following lemma gives the cardinality of the respective factor rings and this proves the theorem. q.e.d.

Lemma II.3.2 *Let R be a principal ideal domain and $a \in R$. Then $\left|R/(a^t)\right| = |R/(a)|^t$.*

Proof. Let $k, l, t \in \mathbb{N}, k \leq l \leq t$ and let $M = (a^t), L = (a^l), K = (a^k)$. Then K, L, M are R-modules and $M \subset L \subset K$ is a chain of submodules. Therefore, the well-known formula $|K/M| = |K/L|\,|L/M|$ gives

$$\begin{aligned}
\left|R/(a^t)\right| &= |R/(a)| \cdot \left|(a)/(a^t)\right| = |R/(a)| \cdot \left|(a)/(a^2)\right| \cdot \left|(a^2)/(a^t)\right| \\
&= \ldots = |R/(a)| \cdot \left|(a)/(a^2)\right| \cdots \left|(a^{t-1})/(a^t)\right| \\
&= |R/(a)|^t .
\end{aligned}$$

q.e.d.

II.3.2.1 Two special cases: rings of integers and polynomial rings

As we have seen above, we are left with counting the number of elements in factor rings of the form $R/(\pi)$, where R is a principal ideal domain and $\pi \in R$ is irreducible. Unfortunately, these rings can have infinitely many elements (for example if $R = K[X]$ for an infinite field K and $\pi = X$). Even if these rings have finitely many elements, it may not be possible to determine the number of elements. In this paragraph we will consider two special cases, namely $R = \mathbb{F}[X]$ where $\mathbb{F}$ is a finite field and $R = \mathcal{O}_K$, i.e., R is the ring of integers of a number field K. We get the following result.

Corollary II.3.3

1. *Let $\mathbb{F}$ be a finite field with exactly p^s elements and $f \in \mathbb{F}[X]$. Let further $m, n \in \mathbb{N}$, $A \in \mathcal{M}_{m,n}(\mathbb{F}[X])$, and $B = \operatorname{diag}(d_1, \ldots, d_r)$ the Smith normal form of A where $r = \min(m, n)$. Then*

$$|\{\boldsymbol{x} \in (\mathbb{F}[X]/(f))^n : A\boldsymbol{x} \in (f)\}| = \prod_{\pi | f} \prod_{i=1}^{r} p^{s \deg(\pi)(\min(v_\pi(d_i), v_\pi(f)))}.$$

2. *Let K be a number field whose ring of integers $\mathcal{O}_K$ is a principal ideal domain (compare Theorem 0.5.26) and let $a \in \mathcal{O}_K$. Let $m, n \in \mathbb{N}$, $A \in \mathcal{M}_{m,n}(\mathcal{O}_K)$, and $B = \operatorname{diag}(d_1, \ldots, d_r)$ be the Smith normal form of A where $r = \min(m, n)$. Then*

$$|\{\boldsymbol{x} \in (\mathcal{O}_K/(a))^n : A\boldsymbol{x} \in (a)\}| = \prod_{\pi | a} \prod_{i=1}^{r} \left| N_{\mathbb{Q}}^{K}(\pi) \right|^{\min(v_\pi(d_i), v_\pi(a))}.$$

3. *Let $a \in \mathbb{Z}$ and $m, n \in \mathbb{N}$. Let further $A \in \mathcal{M}_{m,n}(\mathbb{Z})$ with Smith normal form $B = \operatorname{diag}(d_1, \ldots, d_r)$ where $r = \min(m, n)$. If $m = n$, let $\Delta = |\det(A)|$. Then*

$$|\{\boldsymbol{x} \in (\mathbb{Z}/a\mathbb{Z})^n : A\boldsymbol{x} \equiv \boldsymbol{0} \bmod a\}| = \prod_{p | a} \prod_{i=1}^{r} p^{\min(v_p(d_i), v_p(a))},$$

and if $m = n$ and $\Delta \neq 0$, we have the bound

$$|\{\boldsymbol{x} \in (\mathbb{Z}/a\mathbb{Z})^n : A\boldsymbol{x} \equiv \boldsymbol{0} \bmod a\}| \leq \gcd(\Delta, a^n) \leq \Delta.$$

Proof.

1. First consider $R = \mathbb{F}[X]$ and let $\pi \in R$ be an irreducible polynomial. Then we have
$$|\mathbb{F}[X]/(\pi^w)| = p^{s \deg(\pi) w}.$$
So we get for $f \in \mathbb{F}[X]$
$$\left|\{x_i \in \mathbb{F}[X]/(f) : d_i x_i \in (\pi^{v_\pi(f)})\}\right| = p^{s \deg(\pi)(\min(v_\pi(d_i), v_\pi(f)))}$$
and therefore
$$\left|\{\mathbf{x} \in (\mathbb{F}[X]/(f))^n : B\mathbf{x} \in (\pi^{v_\pi(f)})\}\right| = \prod_{i=1}^{r} p^{s \deg(\pi)(\min(v_\pi(d_i), v_\pi(f)))}.$$
Using the Chinese remainder theorem, this gives
$$|\{\mathbf{x} \in (\mathbb{F}[X]/(f))^n : B\mathbf{x} \in (f)\}| = \prod_{\pi | f} \prod_{i=1}^{r} p^{s \deg(\pi)(\min(v_\pi(d_i), v_\pi(f)))}.$$

2. Let now $R = \mathcal{O}_K$ for a number field K. Then
$$|\mathcal{O}_K/(\pi^w)| = N((\pi^w)) = \left|N_{\mathbb{Q}}^K(\pi^w)\right| = \left|N_{\mathbb{Q}}^K(\pi)\right|^w.$$
For $a \in \mathcal{O}_K$ we obtain
$$\left|\{x_i \in \mathcal{O}_K/(a) : d_i x_i \in (\pi^{v_\pi(a)})\}\right| = \left|N_{\mathbb{Q}}^K(\pi)\right|^{\min(v_\pi(d_i), v_\pi(a))}$$
and therefore
$$\left|\{\mathbf{x} \in (\mathcal{O}_K/(a))^n : B\mathbf{x} \in (\pi^{v_\pi(a)})\}\right| = \prod_{i=1}^{r} \left|N_{\mathbb{Q}}^K(\pi)\right|^{\min(v_\pi(d_i), v_\pi(a))}.$$
Using the Chinese remainder theorem, this gives
$$|\{\mathbf{x} \in (\mathcal{O}_K/(a))^n : B\mathbf{x} \in (a)\}| = \prod_{\pi | a} \prod_{i=1}^{r} \left|N_{\mathbb{Q}}^K(\pi)\right|^{\min(v_\pi(d_i), v_\pi(a))}.$$

3. In the case $K = \mathbb{Q}$ the above formula reads
$$|\{\mathbf{x} \in (\mathbb{Z}/a\mathbb{Z})^n : B\mathbf{x} \equiv \mathbf{0} \bmod a\}| = \prod_{p | a} \prod_{i=1}^{r} p^{\min(v_p(d_i), v_p(a))}.$$

From this equation it is easy to get an upper bound for the number of solutions. If $m = n$ and $\Delta \neq 0$, we get, using $v_p(h_1) + v_p(h_2) = v_p(h_1h_2)$ for all $h_1, h_2 \in R$,

$$\begin{aligned}\prod_{p|a}\prod_{i=1}^{n} p^{\min(v_p(d_i),v_p(a))} &= \prod_p p^{\sum_{i=1}^n \min(v_p(d_i),v_p(a))} \\ &\leq \prod_p p^{\min(v_p(\Delta),nv_p(a))} \\ &= \gcd(\Delta, a^n) \leq \Delta.\end{aligned}$$

q.e.d.

Note that we did not use the condition $d_i|d_{i+1}$. We could use this to give a better upper bound, depending on the index j for which we have $v_p(d_j) < v_p(a)$ and $v_p(d_{j+1}) \geq v_p(a)$ (if such a j exists).

Example II.3.4 Let $A \in \mathcal{M}_{4,4}(\mathbb{Z})$ with $\det(A) = 19440 = 2^4 \cdot 3^5 \cdot 5$. Then there are 30 possibilities for the Smith normal form of A. These possibilities and the corresponding number of solutions of the equation in (II.3.3) with $a = 18$ are shown in Table II.3.1.

d_1	d_2	d_3	d_4	number of solutions		d_1	d_2	d_3	d_4	number of solutions
1	1	1	19440	18		1	1	2	9720	36
1	1	3	6480	54		1	1	4	4860	36
1	1	6	3240	108		1	1	9	2160	162
1	1	12	1620	108		1	1	18	1080	324
1	1	36	540	324		1	2	2	4860	72
1	2	6	1620	216		1	2	18	540	648
1	3	3	2160	162		1	3	6	1080	324
1	3	9	720	486		1	3	12	540	324
1	3	18	360	972		1	3	36	180	972
1	6	6	540	648		1	6	18	180	1944
2	2	2	2430	144		2	2	6	810	432
2	2	18	270	1296		2	6	6	270	1296
2	6	18	90	3888		3	3	3	720	486
3	3	6	360	972		3	3	12	180	972
3	6	6	180	1944		6	6	6	90	3888

Table II.3.1: Possible Smith normal forms of A when $\det(A) = 19440$ and the corresponding number of solutions in (II.3.3) with $a = 18$.

Figure II.3.1 shows the distribution of the number of solutions, i.e., the numbers at the bars indicate how many different Smith normal forms yield this number of solutions.

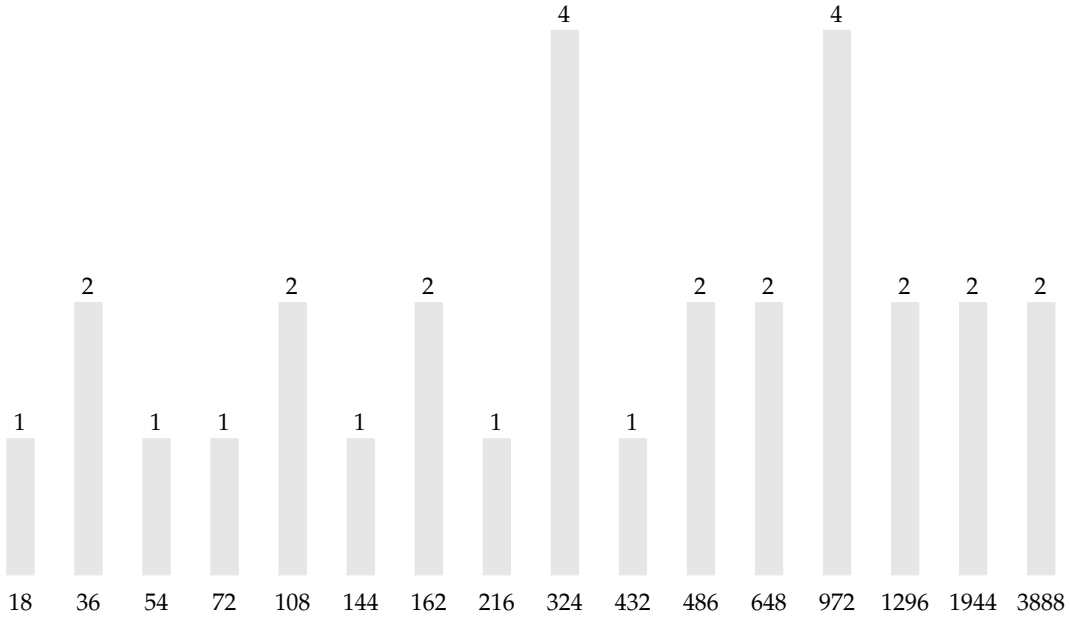

Figure II.3.1: Distribution of the number of solutions of (II.3.3) for matrices with determinant 19440 and $a = 18$.

Table II.3.1 and Figure II.3.1 show that different Smith normal forms can yield the same number of solutions. It would be nice to be able to say something about the distribution of solutions with respect to Smith normal forms (for a given determinant), cf. Section II.3.6.

II.3.3 The number $\left|\{\mathbf{x} \in (R/(a))^n : \langle \mathbf{d}, \mathbf{x} \rangle \in (a)\}\right|$

Now we determine the number

$$|\{\mathbf{x} \in (R/(a))^n : \langle \mathbf{d}, \mathbf{x} \rangle \in (a)\}|$$

directly, i.e., whithout using the notion of Smith normal forms.

As said before, we will impose another condition to this equation. This will be an equivalent of the condition we had in our first problem, i.e., we consider only those $\mathbf{d}$ that satisfy $d_i | d_{i+1}$. Since for $m = n$ each solution of the equation in (II.3.1) is also a solution for the equation in (II.3.2), we see at once that the equation in (II.3.2) has at least as many solutions as the equation in (II.3.1). Precisely, we have

Theorem II.3.5 *Let R be a principal ideal domain such that $|R/(\pi)|$ is finite for each irreducible $\pi \in R$. Let $a \in R$ and $\boldsymbol{d} = (d_1, \ldots, d_n) \in R^n$ with $d_i | d_{i+1}$. Then*

$$|\{\boldsymbol{x} \in (R/(a))^n : \langle \boldsymbol{d}, \boldsymbol{x} \rangle \in (a)\}| = \prod_{\pi | a} |R/(\pi)|^{(n-1)v_\pi(a) + \min(v_\pi(d_1), v_\pi(a))}.$$

Proof. Let $t = v_\pi(a)$. We will again use the Chinese remainder theorem and we first consider the equation $d_1 x_1 + \cdots + d_n x_n \in (\pi^t)$. Choose $x_2, \ldots, x_n \in R/(\pi^t)$. This gives some residue

$$r := \sum_{i=2}^{n} d_i x_i \in R/(\pi^t).$$

Since $v_\pi(d_{i+1}) \geq v_\pi(d_i)$ we have

$$v_\pi(r) \geq v_\pi(d_2) \geq v_\pi(d_1).$$

Let $r = \pi^{v_\pi(r)} f$ and $d_1 = \pi^{v_\pi(d_1)} q$ with $f, q \in (R/(\pi))^*$. We get the equation

$$q\pi^{v_\pi(d_1)} x_1 + \pi^{v_\pi(r)} f \in (\pi^t).$$

If $t \leq v_\pi(d_1)$, then $v_\pi(r) \geq t$ and we can choose x_1 arbitrary in $R/(\pi^t)$. In the other case the equation is equivalent to

$$q x_1 + \pi^{v_\pi(r) - v_\pi(d_1)} f \in \left(\pi^{t - v_\pi(d_1)}\right).$$

We want to multiply this by the inverse of q in $R/\left(\pi^{t-v_\pi(d_1)}\right)$. Since $\pi \nmid q$ we know that q is not a zero divisor in $R/(\pi^{t-v_\pi(d_1)})$. From the well-known fact that in finite rings an element $a \neq 0$ is either a unit or a zero divisor, we deduce that q is a unit. Let $\overline{q}$ be the inverse of q in $R/(\pi^{t-v_\pi(d_1)})$. Then we have

$$x_1 = -f\overline{q}\pi^{v_\pi(r) - v_\pi(d_1)} + \left(\pi^{t - v_\pi(d_1)}\right),$$

i.e., x_1 is uniquely determined in $R/(\pi^{t-v_\pi(d_1)})$. That leaves us with

$$\frac{\left|R/(\pi^t)\right|}{\left|R/\left(\pi^{t-v_\pi(d_1)}\right)\right|} = \left|R/\left(\pi^{v_\pi(d_1)}\right)\right|$$

possible values for x_1. Together with the $\left|R/\left(\pi^t\right)\right|$ choices for each of $x_2, \ldots, x_n$ this proves the theorem. q.e.d.

II.3.3.1 Two special cases: rings of integers and polynomial rings

We compute the number (II.3.2) for the rings $\mathbb{F}[x]$ and $\mathcal{O}_K$.

Corollary II.3.6

1. *Let $\mathbb{F}$ be a finite field with $|\mathbb{F}| = p^s$ and $f \in \mathbb{F}[X]$. Let further $n \in \mathbb{N}$ and $\boldsymbol{d} = (d_1, \dots, d_n) \in (\mathbb{F}[X])^n$ with $d_i | d_{i+1}$ for all $i = 1, \dots, n-1$. Then*

$$|\{\boldsymbol{x} \in (\mathbb{F}[X]/(f))^n : \langle \boldsymbol{d}, \boldsymbol{x} \rangle \in (f)\}| = \prod_{\pi | f} p^{s \deg(\pi)((n-1)v_\pi(f) + \min(v_\pi(f), v_\pi(d_1)))}.$$

2. *Let K be a number field whose ring of integers $\mathcal{O}_K$ is a principal ideal domain (compare Theorem 0.5.26) and let $a \in \mathcal{O}_K$. Let $n \in \mathbb{N}$ and $\boldsymbol{d} = (d_1, \dots, d_n) \in (\mathcal{O}_K)^n$ with $d_i | d_{i+1}$ for all $i = 1, \dots, n-1$. Then*

$$|\{\boldsymbol{x} \in (\mathcal{O}_K/(a))^n : \langle \boldsymbol{d}, \boldsymbol{x} \rangle \in (a)\}| = \prod_{\pi | a} \left| N_{\mathbb{Q}}^K(\pi) \right|^{v_\pi(a)(n-1) + \min(v_\pi(a), v_\pi(d_1))}.$$

3. *Let $a, n \in \mathbb{N}$ and $\boldsymbol{d} = (d_1, \dots, d_n) \in \mathbb{Z}^n$ with $d_i | d_{i+1}$ for all $i = 1, \dots, n-1$. Then we have*

$$|\{\boldsymbol{x} \in (\mathbb{Z}/a\mathbb{Z})^n : \langle \boldsymbol{d}, \boldsymbol{x} \rangle \equiv 0 \bmod a\}| = a^{n-1} \prod_{p | a} p^{\min(v_p(a), v_p(d_1))}$$

and we have the bound

$$|\{\boldsymbol{x} \in (\mathbb{Z}/a\mathbb{Z})^n : \langle \boldsymbol{d}, \boldsymbol{x} \rangle \equiv 0 \bmod a\}| \leq \min(a^n, a^{n-1} d_1).$$

Proof. We have

$$\left| \{\mathbf{x} \in (\mathbb{F}[X]/(f))^n : \langle \mathbf{d}, \mathbf{x} \rangle \in (\pi^{v_\pi(f)})\} \right| = p^{s \deg(\pi) v_\pi(f)(n-1)} p^{s \deg(\pi) \min(v_\pi(f), v_\pi(d_1))}$$

and

$$\left| \{\mathbf{x} \in (\mathcal{O}_K/(a))^n : \langle \mathbf{d}, \mathbf{x} \rangle \in (\pi^{v_\pi(a)})\} \right| = \left| N_{\mathbb{Q}}^K(\pi) \right|^{v_\pi(a)(n-1)} \left| N_{\mathbb{Q}}^K(\pi) \right|^{\min(v_\pi(a), v_\pi(d_1))}.$$

With the Chinese remainder theorem, this gives

$$|\{\mathbf{x} \in (\mathbb{F}[X]/(f))^n : \langle \mathbf{d}, \mathbf{x} \rangle \in (f)\}| = \prod_{\pi | f} p^{s \deg(\pi)((n-1)v_\pi(f) + \min(v_\pi(f), v_\pi(d_1)))}$$

and

$$|\{\mathbf{x} \in (\mathcal{O}_K/(a))^n : \langle \mathbf{d}, \mathbf{x} \rangle \in (a)\}| = \prod_{\pi|a} \left| N_{\mathbb{Q}}^K(\pi) \right|^{v_\pi(a)(n-1)+\min(v_\pi(a),v_\pi(d_1))},$$

and in particular if $K = \mathbb{Q}$, we get

$$\begin{aligned} |\{\mathbf{x} \in (\mathbb{Z}/a\mathbb{Z})^n : \langle \mathbf{d}, \mathbf{x} \rangle \equiv 0 \bmod a\}| &= \prod_{p|a} p^{v_p(a)(n-1)+\min(v_p(a),v_p(d_1))} \\ &= a^{n-1} \prod_{p|a} p^{\min(v_p(a),v_p(d_1))}. \end{aligned}$$

For the upper bound in the case $R = \mathbb{Z}$, we note that

$$a^{n-1} \prod_{p|a} p^{\min(v_p(a),v_p(d_1))} \leq \min(a^n, a^{n-1} d_1).$$

q.e.d.

II.3.4 Remarks

A few remarks shall be made here. As already mentioned, the number (II.3.4) is clearly greater than or equal to the number (II.3.3) if $m = n$. We get this also from our theorems, since

$$\begin{aligned} |\{\mathbf{x} \in (\mathbb{Z}/a\mathbb{Z})^n : A\mathbf{x} \equiv \mathbf{0} \bmod a\}| &= \prod_{p|a} \prod_{i=1}^{n} p^{\min(v_p(d_i),v_p(a))} \\ &\leq \left(\prod_{p|a} \prod_{i=2}^{n} p^{v_p(a)} \right) \cdot \left(\prod_{p|a} p^{\min(v_p(d_1),v_p(a))} \right) \\ &= a^{n-1} \prod_{p|a} p^{\min(v_p(d_1),v_p(a))} \\ &= |\{\mathbf{x} \in (\mathbb{Z}/a\mathbb{Z})^n : \langle \mathbf{d}, \mathbf{x} \rangle \equiv 0 \bmod a\}|. \end{aligned}$$

It was clear from the beginning that the equation in (II.3.4) has at least a^{n-1} solutions. We also see that in the case $v_p(d_2) < v_p(a)$ the equation in (II.3.3) has less solutions then the one in (II.3.4). If we have $v_p(d_i) < v_p(a)$ for many (or even all) indices $i \geq 2$, then the equation in (II.3.3) has much fewer solutions then the one in (II.3.4). This is the reason why we cannot get a good upper bound for (II.3.4).

As a last remark we examine the condition $d_i | d_{i+1}$. While we did not use this in the treatment of (II.3.1), we did so for (II.3.2) when dealing with the residue r. In fact, our arguments would also work if we didn't assume that $d_i | d_{i+1}$, but in the definition of r we would have to distinguish which of the d_i has the smallest π-valuation $v_\pi(d_i)$, for each $\pi | a$ separately. This would make the formula much more complicated.

With the result for (II.3.3) we get the following estimate for the sum $S_q(\mathbf{c})$ mentioned in Section II.3.1:

$$\begin{aligned}
|S_q(\mathbf{c})| &\leq \left(\varphi(q) \sum_{a \bmod q}^{*} \sum_{\mathbf{w} \bmod q} e_q(aF(\mathbf{w}) + \langle \mathbf{c}, \mathbf{w} \rangle) \sum_{\mathbf{v} \bmod q} e_q(\langle a\mathbf{v}, \nabla F(\mathbf{w}) \rangle) \right)^{\frac{1}{2}} \\
&\leq \left(\varphi(q) \sum_{a \bmod q}^{*} q^n \left| \{ \mathbf{v} : (2M)\mathbf{v} \equiv \mathbf{0} \bmod q \} \right| \right)^{\frac{1}{2}} \\
&\leq \left(\varphi(q) \sum_{a \bmod q}^{*} q^n \gcd(2^n \det(M), q) \right)^{\frac{1}{2}} \\
&\leq (\gcd(2^n \det(M), q))^{\frac{1}{2}} q^{1+\frac{n}{2}}.
\end{aligned}$$

II.3.5 Distribution of solutions

Now we examine the distribution of the solutions of the equation in (II.3.3) with respect to the ideal (a). Let A be an $n \times n$ integer matrix, $\Delta := |\det(A)|$, and $m \in \mathbb{N}$. We define

$$\#_y(A, m) := \left| \left\{ a \in [1, m] : \left| \{ \mathbf{x} \in (\mathbb{Z}/a\mathbb{Z})^n : A\mathbf{x} = \mathbf{0} \bmod a \} \right| = y \right\} \right|.$$

We determine $\#_y(A, m)$ in some special cases. First, $\#_y(A, m) = 0$ if $y \nmid \Delta$ (this follows directly from the formula in Theorem II.3.1). The converse is not true, as the examples in Appendix A.4 show. From now on, we suppose that $y | \Delta$. Note that in the case $a = 1$ the equation trivially has exactly one solution.

Theorem II.3.7 *Let $A \in \mathcal{M}_{n,n}(\mathbb{Z})$ and $\Delta = |\det(A)|$. Let $\Delta = \prod_{i=1}^{r} p_i^{v_i}$ be the prime decomposition of Δ.*

1. *For $y = 1$, we have*

$$\#_1(A, m) = S_m^\Delta := \sum_{I \subset \{1,\dots,r\}} (-1)^{|I|} \left\lfloor \frac{m}{\prod_{i \in I} p_i} \right\rfloor$$

and

$$S_m^\Delta = m \prod_{p|\Delta} \left(1 - \frac{1}{p}\right) - R$$

with $R \in (-2^{r-1}, 2^{r-1})$ independently of m.

2. *Let $y > 1$ with $y|\Delta$ and let y, Δ both be square-free. Then we have*

$$\#_y(A, m) = S_m^\Delta(y) := \sum_{I \subset \{1,\dots,r\}} (-1)^{|I|} \left\lfloor \frac{m}{y \cdot \prod_{i \in I} p_i} \right\rfloor$$

and

$$S_m^\Delta(y) = \frac{m}{y} \prod_{\substack{p|\Delta \\ p \nmid y}} \left(1 - \frac{1}{p}\right) - R_y$$

with $R_y \in (-2^{r-1}, 2^{r-1})$ independently of m.

In particular we have $\#_\Delta(A, m) = \lfloor \frac{m}{\Delta} \rfloor$ if Δ is squarefree.

Proof.

1. First we consider the case $y = 1$. We have

$$\left|\{\mathbf{x} \in (\mathbb{Z}/a\mathbb{Z})^n : A\mathbf{x} = \mathbf{0} \bmod a\}\right| = 1$$

if and only if for all p we have

$$v_p(a) = 0 \text{ or } v_p(d_l) = 0 \text{ for all } l,$$

i.e., if we have $p \nmid \Delta$ or $p \nmid a$ for all $p \in \mathbb{P}$. So we get

$$\begin{aligned}
\#_1(A, m) &= |\{a \in [1, m] : (\forall p \in \mathbb{P} : p \nmid \Delta \text{ or } p \nmid a)\}| \\
&= m - |\{a \in [1, m] : (\exists p \in \mathbb{P} : p|\Delta \text{ and } p|a)\}| \\
&= m - |\{a \in [1, m] : p_i|a \text{ for some } i \in \{1, \dots, r\}\}| \\
&= m - \sum_{\emptyset \neq I \subset \{1,\dots,r\}} (-1)^{|I|+1} \left\lfloor \frac{m}{\prod_{i \in I} p_i} \right\rfloor \\
&= S_m^\Delta.
\end{aligned}$$

We estimate S_m^Δ. For all $I \subset \{1,\dots,r\}$ let $\frac{m}{\prod_{i\in I} p_i} = \left\lfloor \frac{m}{\prod_{i\in I} p_i} \right\rfloor + \varepsilon_I$. In particular $0 \le \varepsilon_I < 1$ for all I. Then

$$\begin{aligned}\#_1(A,m) &= \sum_{I\subset\{1,\dots,r\}} (-1)^{|I|} \left\lfloor \frac{m}{\prod_{i\in I} p_i} \right\rfloor \\ &= \sum_{I\subset\{1,\dots,r\}} (-1)^{|I|} \left(\frac{m}{\prod_{i\in I} p_i} + \varepsilon_I \right) \\ &= \sum_{I\subset\{1,\dots,r\}} (-1)^{|I|} \frac{m}{\prod_{i\in I} p_i} + \underbrace{\sum_{I\subset\{1,\dots,r\}} (-1)^{|I|} \varepsilon_I}_{=:R}.\end{aligned}$$

We want to bound R. There are $\binom{r}{k}$ subsets I of $\{1,\dots,r\}$ with exactly k elements. Let $\varepsilon_k \in [0,1)$ be the arithmetic mean of all the ε_I with $|I| = k$. Then

$$R = \sum_{k=0}^{r} (-1)^k \binom{r}{k} \varepsilon_k.$$

Thus we get

$$R > -\sum_{l=0}^{\lfloor \frac{r-1}{2} \rfloor} \binom{r}{2l+1} = -\frac{1}{2}(2^r) > -2^{r-1}$$

and

$$R < \sum_{l=1}^{\lfloor \frac{r}{2} \rfloor} \binom{r}{2l} = \frac{1}{2}(2^r) < 2^{r-1}.$$

2. Let $y = p_1^{f_1} \cdots p_r^{f_r}$ with $f_i \in \{0,1\}$ be the prime decomposition of y. Then

$$\left|\{\mathbf{x} \in (\mathbb{Z}/a\mathbb{Z})^n : A\mathbf{x} = \mathbf{0} \bmod a\}\right| = y$$

if and only if

- $\sum_{l=1}^{r} v_{p_i}(d_l) = 1$ and $v_{p_i}(a) \ge 1$ for all i with $f_i = 1$, and
- $v_{p_i}(d_l) = 0$ or $v_{p_i}(a) = 0$ for all i with $f_i = 0$.

Since Δ is squarefree, the condition $\sum_{l=1}^{r} v_p(d_l) = 1$ is equivalent to $p|\Delta$. Thus we have

$$\left|\{\mathbf{x} \in (\mathbb{Z}/a\mathbb{Z})^n : A\mathbf{x} = \mathbf{0} \bmod a\}\right| = y$$

if and only if

$$(p_i|\Delta \text{ and } p_i|a \text{ if } f_i = 1) \text{ and } (p_i \nmid \Delta \text{ or } p_i \nmid a \text{ if } f_i = 0).$$

In the same way as for $y = 1$ we now get

$$\begin{aligned}\#_y(A,m) &= \sum_{I\subset\{1,\dots,r\}} (-1)^{|I|} \left\lfloor \frac{m}{y\cdot\prod_{i\in I} p_i}\right\rfloor \\ &= S_m^{\Delta}(y) \\ &= \frac{m}{y}\prod_{\substack{p\mid\Delta\\ p\nmid y}}\left(1-\frac{1}{p}\right) - \underbrace{\sum_{k=0}^{r}(-1)^k\binom{r}{k}\varepsilon_k(y)}_{=:R_y}\end{aligned}$$

and we have $R_y \in (-2^{r-1}, 2^{r-1})$ independently of m. This proves the theorem.

q.e.d.

If $y \neq 1$ is not squarefree, the number $\#_y(A,m)$ will depend on the Smith normal form of the matrix A (as can be seen from the examples in Appendix A.4).

II.3.6 Future Work

There are some more questions or generalizations which one could examine. We determined the number of solutions of some linear equations for special rings, this could be done for other special rings.

- Determine the number of solutions of the equations in (II.3.1) and (II.3.2) for other special rings.

It would also be interesting to know if the same results hold if R is not a principal ideal domain but a unique factorization domain.

- Compute the numbers (II.3.1) and (II.3.2) it R is a unique factorization domain.

This may (at least for (II.3.1)) get more complicated, since matrices over unique factorization domains need not have a Smith normal form. In the determination of (II.3.2) we did not need the Smith normal form, thus it could probably be possible to determine (II.3.2) for unique factorization domains. This again justifies the separate treatment of (II.3.2) instead of just handling it as a special case of (II.3.1).

Of course, our approach would not work if we were dealing with nonlinear equations. But one could think about multilinear equations: If $\psi : R^n \to R$ is a multilinear form, one can consider the equation $\psi(\mathbf{x}) \in (a)$. This is a generalization

of (II.3.2) where we are in fact dealing with a linear form ψ, i.e., a linear map $\psi : R \to R$. Any n-linear form can be represented by an n-tensor (this is a matrix if $n = 2$). One could try to determine the corresponding number of solutions:

- Let $\psi : R^n \to R$ be a multilinear form. Determine the number

$$|\{\mathbf{x} \in (R/(a))^n : \psi(\mathbf{x}) \in (a)\}|.$$

This seems to be much harder, already for the case $n = 2$ where we have to deal with matrices. If the bilinear form is represented by the matrix A, we can again try to use the Smith normal form to obtain

$$\mathbf{x}^T A\mathbf{x} = (U^T\mathbf{x})^T A(V\mathbf{x}).$$

Here we cannot substitute $\mathbf{y} = V\mathbf{x}$ to get $\mathbf{y}^T A\mathbf{y}$ (this would be the equivalent to what we did in the linear case). If we substitute to $\mathbf{y(x)}^T A\mathbf{y}$ (i.e., $\mathbf{y}$ depends on $\mathbf{x}$) and write the dependence on $\mathbf{x}$ explicitly, this would destroy the nice form of A. We could of course start with a matrix A in Smith normal form (in fact we did this in (II.3.2)), but for $n \geq 2$ we cannot use the same (but more technical) argument for general A. Thus, one probably has to find another method.

One could also examine the distribution of solutions more closely. On one hand, one could determine $\#_y(A, m)$ for (more) general determinants Δ and values y (not only for the squarefree case). This would (at least in some cases) depend on the Smith normal form of A.

- Determine $\#_y(A, m)$ for (arbitrary) A and y.

On the other hand, one could try to be more exact in the formulae in Section II.3.5, i.e., one could try to find better bounds for R and R_y.

- Find better bounds on R and R_y in Theorem II.3.7 (maybe depending not only on the number of primes dividing Δ, but on A).

One could also examine the distribution of solutions with respect to different matrices rather than different ideals (compare Figure II.3.1):

- Let $m, n, \Delta \in \mathbb{N}$. Consider all $n \times n$ matrices with entries in $\mathbb{Z}/m\mathbb{Z}$ and determinant Δ. How is the number of solutions distributed?

II.4

Lights Out

Here we examine the puzzle "Lights Out". We will model this puzzle with linear algebra and we will see that we only have to deal with determinants of certain matrices. These will be handled with linear algebra, analysis, and various number theoretical tools, making this chapter a connection between those three areas.

The results will be published in [Kre17]. I would like to thank the editor and the referees for their useful comments. I would also like to thank George Schaeffer for the suggestion of Theorem II.4.13 as well as the idea of its proof, Tom Edgar for the suggestion of the proof of Theorem II.4.10, and Carsten Elsner, who suggested the proof for Lemma II.4.12.

II.4.1 Introduction

We start with describing the rules for "Lights Out":

- The player is given a 5×5 square (which we call the **board**) with buttons that can either be illuminated or not. By pressing a button, the button itself as well as all of its (at most 4) neighbours will change their illumination state.
- At the beginning of the game, some of the buttons are illuminated and some are not. The goal is to get all the buttons turned off by pressing some combination of the buttons.

Example II.4.1 Suppose we begin with the board shown in Figure II.4.1 (here a circle indicates that the button is illuminated). The steps shown in Figure II.4.1 lead to the desired goal, where $\xrightarrow{(a,b)}$ means that we press the button in row a and column b.

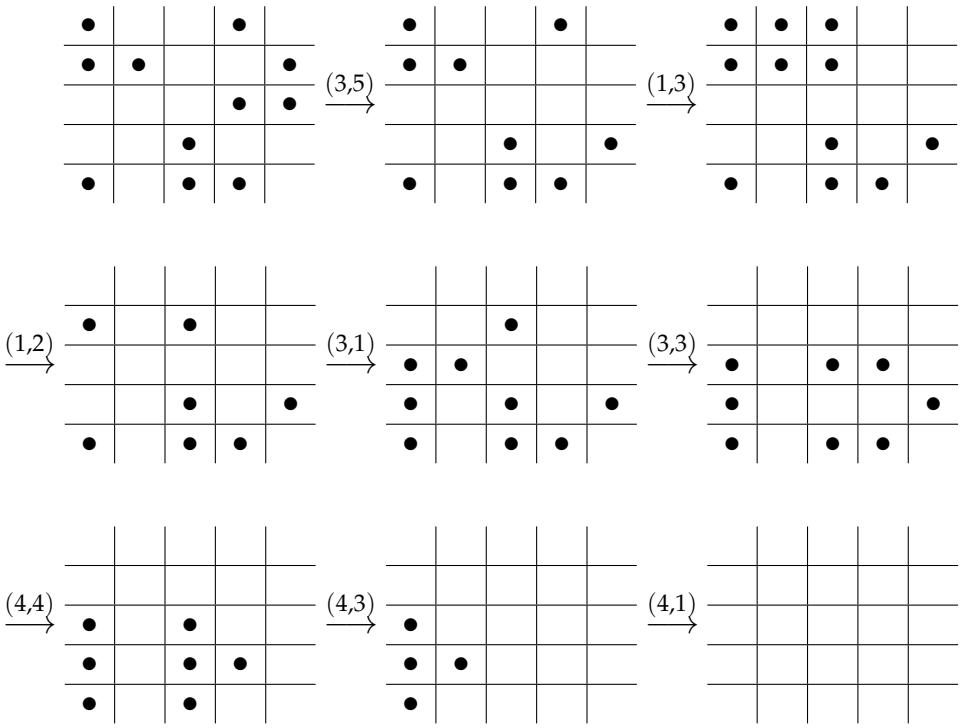

Figure II.4.1: Solving a given "Lights Out" starting board.

There are many variants of the original puzzle (some of which we will address in Section II.4.6). In the main part of this chapter we will deal with a square game board of size $n \times n$, where each button can be lit with k different colors. Here we also assign a color to unlit buttons (thus the original puzzle has exactly two colors) and we always assume that there are at least two colors. When there are more than two colors, the colors will change cyclicly (and for every button in the same way). We will refer to this puzzle as $BLO(n, k)$ (here the "B" stands for "border" in order to differentiate this puzzle from an unbounded variant considered in Section II.4.6). For obvious reasons, we will always let $n \geq 2$.

Example II.4.2 We consider the puzzle $BLO(2, 3)$. We will denote the three colors with $0, 1$, and 2. We define the changing of colors as follows:

$$0 \to 1 \to 2 \to 0$$

So a possible position is

$$\begin{array}{c|c} 2 & 1 \\ \hline 0 & 1 \end{array}$$

and pressing the leftmost button in the first row will result in the board

$$\begin{array}{c|c} 0 & 2 \\ \hline 1 & 1 \end{array}.$$

We are interested in determining if the puzzle is solvable for all starting positions. In this case we call $BLO(n,k)$ **completely solvable**. As we will see, not every puzzle is completely solvable. Even in the original puzzle, i.e., $BLO(5,2)$, there are starting positions in which the puzzle cannot be solved, see [AF98].

Example II.4.3 The puzzle $BLO(2,3)$ is not completely solvable (in fact the starting board in Example II.4.2 cannot be solved). This can be shown very easily, but since the argument is the same as in Proposition II.4.22, we will omit it here.

Further information about "Lights Out" and some variants can be found in [Sin, 7.M.6] and [Schb]. In this chapter, we will only focus on the complete solvability of $BLO(n,k)$. We will neither discuss strategies of winning a game (when the starting position is solvable) nor investigate which starting boards are not solvable. Winning strategies will arise directly from the modeling of the problem. More direct strategies (such as "Light Chasing") are discussed in [Scha, Schb, Sol]. The unsolvable starting boards for the original "Lights Out" game $BLO(5,2)$ have been characterized in [AF98]. A great overview of the games and variants, as well as some first mathematical results can be found in [Schb].

We will model the problem with linear algebra (but will need some number theory and analysis in later proofs). There are other ways to examine this problem as well as variants of it. Fibonacci polynomials have been used in [GKT97, Klo] to study the complete solvability of games on an $m \times n$ rectangular board (but with only two colors). There are numerous papers that investigate light flipping games on graphs, which are generalizations of the original "Lights Out", see [BR96, CG, DW01, EES01, GK97, Sut89]. Another generalization of "Lights Out" is known as "Confused Electrician Games". In [ES16], the authors investigate such games using the Smith normal form. A similar game is "Berlekamp's switching game", which has been studied in [FS89], with an error that was corrected in [CS04].

The outline of this chapter is as follows: In Section II.4.2 we will model the problem with linear algebra. We will see that we only have to deal with determinants of certain matrices afterwards. In Section II.4.3 we will determine the cases

in which this determinant is 0 and thus determine the cases in which $BLO(n,k)$ (for given n) is not completely solvable for any k. In Section II.4.4 we show that there is no n such that $BLO(n,k)$ is completely solvable for all k. We use another approach to examine the complete solvability, relying on algebraic number theory, in Section II.4.5. Section II.4.6 contains a variant of the original puzzle. We will see that this variant is easier to handle and we show the corresponding results.

II.4.2 Modeling the problem

We will view the $n \times n$ board of the puzzle $BLO(n,k)$ as an $n \times n$ matrix. Further, we will enumerate the colors by $0, \ldots, k-1$ according to the cyclic changing of the colors, where 0 means that the light is out. Then, at any time, the board can be viewed as a matrix $A \in \mathcal{M}_{n,n}(\mathbb{Z}/k\mathbb{Z})$. Pressing a button (i,j) will result in adding a certain matrix $B_{i,j}$ to A. The added matrix $B_{i,j}$ has a 1 at the place where a color will be changed and a 0 otherwise, e.g., for the puzzle $BLO(5,k)$ we have

$$B_{1,1} = \begin{pmatrix} 1 & 1 & & & \\ 1 & & & & \\ & & & & \\ & & & & \\ & & & & \end{pmatrix}, \quad B_{1,4} = \begin{pmatrix} & & & 1 & 1 \\ & & & & 1 \\ & & & & \\ & & & & \\ & & & & \end{pmatrix}, \quad B_{4,2} = \begin{pmatrix} & & & & \\ & & & & \\ & 1 & & & \\ 1 & 1 & 1 & & \\ & 1 & & & \end{pmatrix}.$$

We will denote the zero matrix by $\mathbf{0}$. Then, given a starting board A, we need to find $c_{i,j} \in \mathbb{Z}/k\mathbb{Z}$ such that

$$A + \sum_{i,j=1}^{n} c_{i,j} B_{i,j} = \mathbf{0},$$

i.e.,

$$\sum_{i,j=1}^{n} c_{i,j} B_{i,j} = -A. \tag{II.4.1}$$

If we consider each entry of the matrices involved individually, we obtain a system of n^2 linear equations. The corresponding matrix (when going systematically from left to right and top to bottom) has the following form.

$$BL_n := \begin{pmatrix} J_n & I_n & 0 & \cdots & 0 \\ I_n & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & I_n \\ 0 & \cdots & 0 & I_n & J_n \end{pmatrix}$$

where I_n is the $n \times n$ identity matrix and

$$J_n := \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 1 & 1 \end{pmatrix}, \text{ i.e., } (J_n)_{i,j} = \begin{cases} 1, & |i-j| \in \{0,1\} \\ 0, & \text{otherwise} \end{cases}.$$

The matrix J_n is a **Toeplitz tridiagonal** matrix. These have been widely studied, see [NPR13, Yue05]. We will only need the eigenvalues for our special case. These are

$$\lambda_j = 1 + 2\cos\left(\frac{j\pi}{n+1}\right), \text{ where } 1 \leq j \leq n. \tag{II.4.2}$$

II.4.3 (Un)Solvability of Lights Out

Obviously, $BLO(n,k)$ is completely solvable if and only if System (II.4.1) is solvable for all A. But this is equivalent to saying that BL_n is invertible over $\mathbb{Z}/k\mathbb{Z}$. Thus we obtain the following two results.

Proposition II.4.4 *Let n and k be natural numbers and let $n \geq 2$.*

1. *The game $BLO(n,k)$ is completely solvable if and only if $\det(BL_n)$ is invertible modulo k, i.e., if and only if $\gcd(\det(BL_n), k) = 1$.*

2. *Fix n. Then $BLO(n,k)$ is completely solvable for no k if and only if $\det(BL_n) = 0$, and $BLO(n,k)$ is completely solvable for all k if and only if $\det(BL_n) = \pm 1$.*

Lemma II.4.5 *$BLO(n,k)$ is completely solvable if and only if $BLO(n,p)$ is completely solvable for all $p \in \mathbb{P}$ with $p|k$.*

Proof. This follows directly from the fact that $BLO(n,k)$ is completely solvable if and only if $\gcd(\det(BL_n), k) = 1$. q.e.d.

This lemma will not be needed for computing the determinant in the general case. However, we will explicitly compute the deteterminant for some small n and, in view of this lemma, we only need to know the prime factors of the determinant.

Now we determine $\det(BL_n)$. The arguments below follow [Pav] and [Jam]. First note that we can write

$$BL_n = J_n \otimes I_n + I_n \otimes J_n - I_{n^2}$$

where $\otimes$ denotes the Kronecker product. According to (II.4.2) and the properties of the Kronecker product mentioned in Section 0.4, the eigenvalues of $J_n \otimes I_n$ are $1 + 2\cos\left(\frac{j\pi}{n+1}\right)$, and the eigenvalues of $I_n \otimes J_n$ are $1 + 2\cos\left(\frac{l\pi}{n+1}\right)$. Since the matrices J_n and I_n commute, the matrices $J_n \otimes I_n$ and $I_n \otimes J_n$ commute. Therefore, the eigenvalues of BL_n are

$$\lambda_{j,l} = 1 + 2\left(\cos\left(\frac{j\pi}{n+1}\right) + \cos\left(\frac{l\pi}{n+1}\right)\right)$$

and we have

$$\det(BL_n) = \prod_{j=1}^{n}\prod_{l=1}^{n}\lambda_{j,l}.$$

Tables II.4.1 and II.4.2 show the nonzero determinants of BL_n for $2 \leq n \leq 30$, respectively, the prime decomposition of its absolute value (computed with Mathematica).

n	$\det(BL_n)$	prime decomposition of $\lvert\det(BL_n)\rvert$
2	-3	3^1
3	-7	7^1
6	2197	13^3
7	-34391	$7^1 \cdot 17^3$
8	-4002939	$3^6 \cdot 17^2 \cdot 19^1$
10	276762749	$23^5 \cdot 43^1$
12	-133968364171875	$3^6 \cdot 5^6 \cdot 53^3 \cdot 79^1$
13	-239121867667810023	$3^6 \cdot 13^3 \cdot 29^6 \cdot 251^1$
15	$1054995627763436597175 77$	$7^1 \cdot 17^3 \cdot 31^8 \cdot 127^2 \cdot 223^1$
16	$-3916466797684156666150912$	$2^{16} \cdot 67^5 \cdot 101^2 \cdot 103^2 \cdot 409^1$
18	$-6847054013331284711313441844382 51$	$37^{11} \cdot 113^4 \cdot 191^2 \cdot 647^1$

Table II.4.1: The nonzero determinants $\det(BL_n)$ for $2 \leq n \leq 19$ and their prime decompositions.

n	prime decomposition of $\lvert\det(BL_n)\rvert$
20	$3^1 \cdot 7^8 \cdot 13^3 \cdot 41^7 \cdot 43^8 \cdot 83^2 \cdot 379^1$
21	$23^5 \cdot 43^9 \cdot 89^6 \cdot 131^2 \cdot 263^2 \cdot 1451^1$
22	$47^{10} \cdot 137^4 \cdot 139^4 \cdot 277^2 \cdot 919^2 \cdot 1747^1$
25	$3^6 \cdot 5^6 \cdot 53^{15} \cdot 79^1 \cdot 103^6 \cdot 157^4 \cdot 727^2 \cdot 2339^1$
26	$3^{23} \cdot 17^2 \cdot 19^1 \cdot 53^{14} \cdot 107^8 \cdot 109^2 \cdot 269^1 \cdot 379^2 \cdot 431^1$
27	$3^6 \cdot 7^{18} \cdot 13^5 \cdot 29^6 \cdot 113^2 \cdot 167^6 \cdot 223^2 \cdot 251^1 \cdot 281^2 \cdot 449^2 \cdot 617^1 \cdot 3527^2$
28	$17^2 \cdot 59^{14} \cdot 173^6 \cdot 233^4 \cdot 347^4 \cdot 463^2 \cdot 5279^2 \cdot 6959^1 \cdot 20011^1$
30	$2^{40} \cdot 5^{12} \cdot 61^{16} \cdot 311^2 \cdot 373^2 \cdot 433^1 \cdot 619^4 \cdot 929^2 \cdot 6449^2 \cdot 53507^1$

Table II.4.2: The prime decomposition of the nonzero determinants $\lvert\det(BL_n)\rvert$ for $20 \leq n \leq 30$.

In view of Proposition II.4.4, we want to know for which values of n the corresponding determinant will be 0 or ± 1. The first case is equivalent to saying that $\lambda_{j,l} = 0$ for some j, l.

We will use the following result.

Theorem II.4.6 *[CJ76, Theorem 7] Suppose we have at most four distinct rational multiples of π lying strictly between 0 and $\pi/2$ for which some rational linear combination of their cosines is rational but no proper subset has this property. Then the appropriate linear combination is proportional to one from the following list:*

$$\begin{aligned}
\cos\left(\frac{\pi}{3}\right) &= \frac{1}{2}, \\
-\cos(\varphi) + \cos\left(\frac{\pi}{3} - \varphi\right) + \cos\left(\frac{\pi}{3} + \varphi\right) &= 0 \qquad \left(0 < \varphi < \frac{\pi}{6}\right), \\
\cos\left(\frac{\pi}{5}\right) - \cos\left(\frac{2\pi}{5}\right) &= \frac{1}{2}, \\
\cos\left(\frac{\pi}{7}\right) - \cos\left(\frac{2\pi}{7}\right) + \cos\left(\frac{3\pi}{7}\right) &= \frac{1}{2}, \\
\cos\left(\frac{\pi}{5}\right) - \cos\left(\frac{2\pi}{15}\right) + \cos\left(\frac{4\pi}{15}\right) &= \frac{1}{2}, \\
-\cos\left(\frac{2\pi}{5}\right) + \cos\left(\frac{2\pi}{15}\right) - \cos\left(\frac{7\pi}{15}\right) &= \frac{1}{2}, \\
\cos\left(\frac{\pi}{7}\right) + \cos\left(\frac{3\pi}{7}\right) - \cos\left(\frac{\pi}{21}\right) + \cos\left(\frac{8\pi}{21}\right) &= \frac{1}{2}, \\
\cos\left(\frac{\pi}{7}\right) - \cos\left(\frac{2\pi}{7}\right) + \cos\left(\frac{2\pi}{21}\right) - \cos\left(\frac{5\pi}{21}\right) &= \frac{1}{2}, \\
-\cos\left(\frac{2\pi}{7}\right) + \cos\left(\frac{3\pi}{7}\right) + \cos\left(\frac{4\pi}{21}\right) + \cos\left(\frac{10\pi}{21}\right) &= \frac{1}{2}, \\
-\cos\left(\frac{\pi}{15}\right) + \cos\left(\frac{2\pi}{15}\right) + \cos\left(\frac{4\pi}{15}\right) - \cos\left(\frac{7\pi}{15}\right) &= \frac{1}{2}.
\end{aligned}$$

Theorem II.4.7 *BL_n is not invertible if and only if $n+1$ is a multiple of 5 or 6.*

Proof. For $n = 4$ we have

$$\begin{aligned}\lambda_{2,4} &= 1 + 2\left(\cos\left(\frac{2\pi}{5}\right) + \cos\left(\frac{4\pi}{5}\right)\right)\\ &= 1 + 2\left(\frac{1}{4}\left(-1+\sqrt{5}\right) + \frac{1}{4}\left(-1-\sqrt{5}\right)\right) = 0,\end{aligned}$$

and for $n = 5$ we have

$$\lambda_{3,4} = 1 + 2\left(\cos\left(\frac{\pi}{2}\right) + \cos\left(\frac{2\pi}{3}\right)\right) = 0,$$

so we get $\det(BL_4) = \det(BL_5) = 0$. Now if $n+1$ is a multiple of 5 or 6, we can choose j, l such that the eigenvalues give the values above. Thus $\det(BL_n) = 0$ if $n+1$ is a multiple of 5 or 6.

We need to show that in the other cases 0 is not an eigenvalue. Suppose that

$$1 + 2\left(\cos\left(\frac{j\pi}{n+1}\right) + \cos\left(\frac{l\pi}{n+1}\right)\right) = 0$$

for some j, l, n. Then $\cos\left(\frac{j\pi}{n+1}\right) + \cos\left(\frac{l\pi}{n+1}\right) \in \mathbb{Q}$. But then $\cos(\alpha\pi) + \cos(\beta\pi)$ is rational for some $\alpha, \beta \in \mathbb{Q}, 0 \leq \alpha, \beta \leq \frac{1}{2}$. Using Theorem II.4.6 we either have $\alpha = \frac{1}{3}, \beta \in [0, \frac{1}{2}]$ or $\alpha = \frac{1}{5}, \beta = \frac{2}{5}$, i.e., $n+1$ has to be a multiple of 3 or 5. If $n+1$ is a multiple of 3, then we need some $\beta \in \mathbb{Q}, \beta \in (0,1)$ such that either

$$1 + 2\left(\cos\left(\frac{\pi}{3}\right) + \cos(\beta\pi)\right) = 2 + 2\cos(\beta\pi) = 0$$

or

$$1 + 2\left(\cos\left(\frac{2\pi}{3}\right) + \cos(\beta\pi)\right) = 2\cos(\beta\pi) = 0.$$

(Here we can take $\beta \in (0,1)$ since we now allow all possible values $\beta = \frac{l}{n+1}$, not only those appearing in Theorem II.4.6.)

The first case is impossible, since this holds only for $\beta \in 2\mathbb{Z}+1$. The second case holds if and only if $\beta = \frac{2l+1}{2}$ for some $l \in \mathbb{Z}$. Thus, $n+1$ has to be a multiple of 2 and therefore a multiple of 6. This proves the theorem. q.e.d.

The statement of Theorem II.4.7 is equivalent to saying that the adjacency matrix of the $n \times n$ grid has eigenvalue -1, see the OEIS entry A162698 [OEI]. Here the $n \times n$ **grid** is the graph $G = (V, E)$ with $V = \{(a,b) : 1 \leq a, b \leq n\}$ and $\{(a,b),(c,d)\} \in E$ if and only if

- $a = c$ and $|b - d| = 1$ or
- $b = d$ and $|a - c| = 1$,

see Figure II.4.2 for $n = 4$.

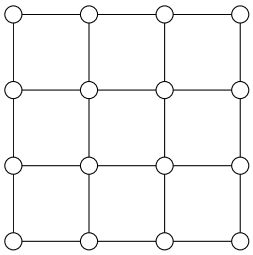

Figure II.4.2: The 4×4 grid.

As corollary of Theorem II.4.7 we get

Corollary II.4.8 *Let $n + 1$ be a multiple of 5 or 6 and k arbitrary. Then $BLO(n, k)$ is not completely solvable.*

This follows immediately from the modeling of "Lights Out". Corollary II.4.8 also gives a nice result for greatest common divisors of Fibonacci polynomials $f_n(x)$, defined by the recurrence

$$f_{n+1}(x) = xf_n(x) + f_{n-1}(x), \quad f_1(x) = 1, \quad f_2(x) = x.$$

Corollary II.4.9 *Let $n + 1$ be a multiple of 5 or 6. Then $f_{n+1}(x)$ and $f_{n+1}(x + 1)$ have a common factor over $\mathbb{Z}/2\mathbb{Z}$.*

Proof. Note that from [GKT97, Theorem 3] we know that $BLO(n, 2)$ is completely solvable if and only if $f_{n+1}(x)$ and $f_{n+1}(x + 1)$ are coprime over $\mathbb{Z}/2\mathbb{Z}$. Thus the claim follows from Corollary II.4.8. q.e.d.

II.4.4 On complete solvability for all k

In this section we will show that, for given n that is neither a multiple of 5 nor a multiple of 6, $BLO(n, k)$ cannot be solvable for all k. In fact, this has first been a conjecture which I was unable to prove. There were quite a few ideas to prove this, such as using lower bounds for determinants (see, for example, [KP01]), formulae for determinants of block matrices (see, for example, [KSW99, Mol08, Tis87]), or the Smith normal form to investigate elementary divisors. One could also deal with the eigenvalues $\lambda_{j,l}$ as elements in some ring (compare Section II.4.5) and try to show that one of these eigenvalues is a non-unit.

None of these methods worked, but the desired result can be proved with different number theoretic tools. In fact, the proofs of both theorems in this section are (at least partially) due to others: The proof of Theorem II.4.10 is mostly due to Tom Edgar, while the idea of the proof of Theorem II.4.13 is due to George Schaeffer.

Theorem II.4.10 *We have* $\det(BL_n) \neq \pm 1$ *for* $n \geq 2$.

Proof. Let

$$\Lambda_n = \prod_{j=1}^{n} \left(1 + 4\cos\left(\frac{j\pi}{n+1}\right)\right),$$
$$V_n = \prod_{\substack{j,l=1 \\ j\neq l}}^{n} \left(1 + 2\cos\left(\frac{j\pi}{n+1}\right) + 2\cos\left(\frac{l\pi}{n+1}\right)\right).$$

Then $\det(BL_n) = \Lambda_n \cdot V_n$. We show that $\Lambda_n, V_n \in \mathbb{Z}$. Since Λ_n is the determinant of the integer matrix

$$\begin{pmatrix} 1 & 2 & 0 & \cdots & \cdots & 0 \\ 2 & 1 & 2 & \ddots & & \vdots \\ 0 & 2 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 2 & 0 \\ \vdots & & \ddots & 2 & 1 & 2 \\ 0 & \dots & \cdots & 0 & 2 & 1 \end{pmatrix}$$

(see [Yue05]), we immediately get $\Lambda_n \in \mathbb{Z}$. Regarding V_n, note that $\Lambda_n \neq 0$ due to Theorem II.4.6, thus $V_n = \frac{\det(BL_n)}{\Lambda_n} \in \mathbb{Q}$. Let $\zeta_{2(n+1)}$ be a primitive $(2n+2)$-nd root of unity. Then $\lambda_{j,l} \in \mathcal{O}_K$ for $K = \mathbb{Q}(\zeta_{2(n+1)})$ (compare Remark II.4.14), thus we get $V_n \in \mathcal{O}_K \cap \mathbb{Q} = \mathbb{Z}$, as required.

Hence Λ_n is an integer dividing $\det(BL_n)$, so it suffices to show that $\Lambda_n \neq \pm 1$. Let (u_m) be the recursive sequence defined by

$$u_m = u_{m-1} - 4u_{m-2}, \quad u_0 = 0, \quad u_1 = 1.$$

The sequence (u_m) is a **Lucas sequence** and we have $\Lambda_n = u_{n+1}$, see [Luc78, Section IV]. Let $a := \frac{1+\sqrt{-15}}{2}$ and $b := \frac{1-\sqrt{-15}}{2}$. Then we have

$$u_m = \frac{a^m - b^m}{a-b}, \qquad a+b = 1, \quad ab = 4, \quad \frac{a}{b} = -\frac{7}{8} + \frac{\sqrt{15}}{8}i.$$

Due to Theorem II.4.6, $\arccos\left(-\frac{7}{8}\right)$ is not a rational multiple of π, hence $\frac{a}{b}$ is not a root of unity. Thus a result of Bilu, Hanrot, and Voutier [BHV01] shows that for all $m > 30$ the integer u_m has a prime divisor p (with an additional property that is not needed here). Hence $\Lambda_n \neq \pm 1$ for $n \geq 30$ and thus we have $\det(BL_n) \neq \pm 1$ for $n \geq 30$. Together with the values of $\det(BL_n)$ in Tables II.4.1 and II.4.2 this proves the theorem. q.e.d.

For the puzzle $BLO(n,k)$ we obtain the following.

Corollary II.4.11 *If $n+1$ is neither a multiple of 5 nor a multiple of 6, there is a k_1 such that $BLO(n,k_1)$ is completely solvable and a k_2 such that $BLO(n,k_2)$ is not completely solvable.*

There is a weaker version of Theorem II.4.10, namely Theorem II.4.13, that states that $\det(BL_n) \neq \pm 1$ holds for all big enough n. Of course, this result is obsolete due to Theorem II.4.10. We include this theorem nevertheless for a particular reason: Its proof uses Diophantine approximation and convergence of Riemann sums to improper integrals. Thus this proof is a connection between number theory and analysis and hence exactly in the spirit of this thesis.

We begin with a lemma about Diophantine approximation.

Lemma II.4.12 (due to C. Elsner) *Let $\xi = \frac{1}{\pi}\arccos\left(-\frac{1}{4}\right)$. There is a constant δ such that for any $a, b \in \mathbb{N}$ with $a < b$ and $b \geq 2$ we have*

$$\left|\xi - \frac{a}{b}\right| > \frac{1}{\pi b^{\delta}}.$$

Proof. Note that $\pi = -i\log(-1)$ and $\arccos(z) = -i\log(z + i\sqrt{1-z^2})$ for $z \in \mathbb{C}$. (Here we take the principal value of the complex logarithm. For the latter identity see entry **1.622** (2) in [GR15].) With $\alpha_1 = \frac{-1+i\sqrt{15}}{4}$ and $\alpha_2 = -1$ we thus get $\xi = \frac{\log(\alpha_1)}{\log(\alpha_2)}$. For $a, b \in \mathbb{N}$ with $a < b$ and $b \geq 2$ let L be the linear form in logarithms

$$L := b\log(\alpha_1) - a\log(\alpha_2).$$

Since α_1 has minimal polynomial $2z^2 + z + 2$, it is algebraic of degree 2 and has height 2. Further, α_1 has degree and height 1, and a and b have degree 1. Since ξ is irrational (this follows from Theorem II.4.6), we have $L \neq 0$ and hence Baker's theorem on linear forms in logarithms (Theorem 0.5.21) yields

$$|L| > b^{-\eta},$$

where η depends only on $n = 2, d = 2$, and $h = 2$, i.e., η is a constant. Dividing the above inequality by $b\,|\log(\alpha_2)| = b\pi$ gives

$$\left|\xi - \frac{a}{b}\right| > \frac{1}{\pi b^{\delta}}$$

with $\delta = \eta + 1$. q.e.d.

Theorem II.4.13 *There is an $N \in \mathbb{N}$ such that $\det(BL_n) \neq \pm 1$ for all $n \geq N$.*

Proof. Define Λ_n and V_n as in the proof of Theorem II.4.10, i.e.,

$$\Lambda_n = \prod_{j=1}^{n} \left(1 + 4\cos\left(\frac{j\pi}{n+1}\right)\right),$$

$$V_n = \prod_{\substack{j,l=1 \\ j\neq l}}^{n} \left(1 + 2\cos\left(\frac{j\pi}{n+1}\right) + 2\cos\left(\frac{l\pi}{n+1}\right)\right),$$

and let further $A_n := |\Lambda_n|$. We show that $A_n > 1$ for $n \geq N$. Consider the sum

$$\frac{1}{n+1}\sum_{j=1}^{n} \log\left|1 + 4\cos\left(\frac{j\pi}{n+1}\right)\right| = \frac{\log A_n}{n+1}.$$

Note that

$$\Omega_{n+1} := \frac{1}{n+1}\sum_{j=1}^{n+1} \log\left|1 + 4\cos\left(\frac{j\pi}{n+1}\right)\right|$$

is a right-handed Riemann sum for the improper integral $\int_0^1 \log|1 + 4\cos(\pi x)|\,\mathrm{d}x$ and

$$\frac{\log A_n}{n+1} = \Omega_{n+1} - \frac{\log 3}{n+1} \to \Omega_{n+1} \text{ for } n \to \infty.$$

We will show that Ω_n converges to the value of $\int_0^1 \log|1 + 4\cos(\pi x)|\,\mathrm{d}x$ and that this value is positive.

For the value of the integral, we get

$$\int_0^1 \log|1 + 4\cos(\pi x)|\,\mathrm{d}x = \log 2.$$

This is entry 172 in [Zwi05], as correction of entry **4.224** (12) in [GR15] which is erroneous.

The integrand $f(x) := \log|1+4\cos(\pi x)|$ is defined for all $x \in [0,1]$ except for $\xi = \frac{1}{\pi}\arccos\left(-\frac{1}{4}\right) \approx 0.580431$. According to Theorem II.4.6, ξ is irrational, so ξ will never be one of the points $\frac{j}{n}$. Choose j_n such that $\frac{j_n}{n} < \xi < \frac{j_n+1}{n}$ and let

$$R_L(n) = \frac{1}{n}\sum_{k=1}^{j_n} f\left(\frac{k}{n}\right), \quad R_R(n) = \frac{1}{n}\sum_{k=j_n+2}^{n} f\left(\frac{k}{n}\right), \quad R_M(n) = \frac{1}{n} f\left(\frac{j_n+1}{n}\right),$$

and

$$I_1 = \int_0^{\xi} f(x)\,\mathrm{d}x, \quad I_2 = \int_{\xi}^{1} f(x)\,\mathrm{d}x.$$

Since the value of the integral $\int_0^1 f(x)\,\mathrm{d}x$ is finite and $f(x)$ is bounded from above, the improper integrals I_1 and I_2 exist. Hence we have

$$\int_0^1 f(x)\,\mathrm{d}x = I_1 + I_2, \qquad \Omega_n = R_L(n) + R_M(n) + R_R(n).$$

Figure II.4.3 shows the function $f(x)$ and the sums $R_L(20)$, $R_M(20)$, and $R_R(20)$.

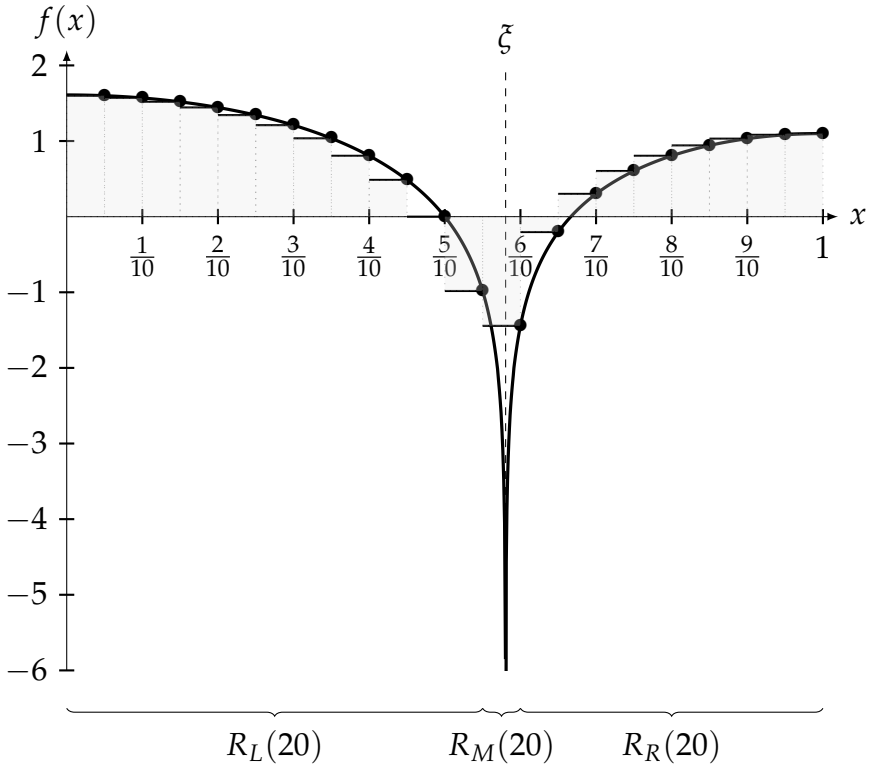

Figure II.4.3: The function $f(x) = \log|1+4\cos(\pi x)|$ and the sums $R_L(20)$, $R_R(20)$, and $R_M(20)$.

We will show that $R_L(n) \to I_1$, $R_R(n) \to I_2$, and $R_M(n) \to 0$ as $n \to \infty$.

We begin with the convergence of $R_M(n)$. Let $B := \frac{j_n+1}{n} - \xi$. The angle sum identity for cosine and the identity $\sin(\arccos(x)) = \sqrt{1-x^2}$ yield

$$\begin{aligned} R_M(n) &= \frac{1}{n} \log \left| 1 + 4\cos\left(\frac{(j_n+1)\pi}{n}\right)\right| \\ &= \frac{1}{n} \log |1 + 4\cos(\xi\pi + B\pi)| \\ &= \frac{1}{n} \log \left| 1 - \cos(B\pi) - \sqrt{15}\sin(B\pi)\right|. \end{aligned}$$

Since $B \to 0$ for $n \to \infty$ and all involved functions are continuous, we can use the asymptotic expansions $\cos(B) = 1 + \mathcal{O}(B^2)$ and $\sin(B) = B + \mathcal{O}(B^3)$ to get

$$\begin{aligned} R_M(n) &= \frac{1}{n} \log \left| -\sqrt{15}B\pi + \mathcal{O}(B^2)\right| = \frac{1}{n}\left(\log\left(\sqrt{15}B\pi\right) + \log\left(1 + \mathcal{O}(B)\right)\right) \\ &= \frac{1}{n}\log(B) + \mathcal{O}\left(\frac{1}{n}\right). \end{aligned}$$

Due to Lemma II.4.12 we have $B > \frac{1}{\pi n^\delta}$ for some constant δ, i.e.,

$$|R_M(n)| < \left|\frac{1}{n}\log\left(\frac{1}{\pi n^\delta}\right) + \mathcal{O}\left(\frac{1}{n}\right)\right| \to 0.$$

For the convergence of $R_L(n)$ we note that $f(x)$ is monotone decreasing for $x \in [0, \xi)$, thus we have

$$\frac{1}{n} f\left(\frac{k}{n}\right) \geq \int_{\frac{k}{n}}^{\frac{k+1}{n}} f(x)\,\mathrm{d}x \geq \frac{1}{n} f\left(\frac{k+1}{n}\right).$$

Summing this for $k = 0, \ldots, j_n - 1$ yields

$$\frac{1}{n}\sum_{k=0}^{j_n-1} f\left(\frac{k}{n}\right) \geq \int_0^{\frac{j_n}{n}} f(x)\,\mathrm{d}x \geq \frac{1}{n}\sum_{k=1}^{j_n} f\left(\frac{k}{n}\right).$$

Hence we have

$$\int_0^{\frac{j_n}{n}} f(x)\,\mathrm{d}x + \frac{1}{n}f(0) \geq \frac{1}{n}\sum_{k=0}^{j_n} f\left(\frac{k}{n}\right) \geq \int_0^{\frac{j_n}{n}} f(x)\,\mathrm{d}x + \frac{1}{n} f\left(\frac{j_n}{n}\right).$$

Since I_1 exists as improper Riemann integral, we have $\int_0^{\frac{j_n}{n}} f(x)\,\mathrm{d}x \to I_1$ as $n \to \infty$. Analogous to the convergence of $R_M(n)$ we get that $\frac{1}{n} f\left(\frac{j_n}{n}\right)$ converges to 0. Using $\frac{1}{n}f(0) \to 0$ we get with the squeeze theorem

$$\lim_{n\to\infty} R_L(n) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{j_n} f\left(\frac{k}{n}\right) = I_1.$$

Since $f(x)$ is monotone increasing for $x \in (\xi, 1]$, the convergence of $R_R(n)$ to I_2 is analogous to that of $R_L(n)$ to I_1.

In conclusion, Ω_n (and thus $\frac{\log A_n}{n+1}$) converges to $\int_0^1 \log|1 + 4\cos(\pi x)|\,\mathrm{d}x$. Since

$$\int_0^1 \log|1 + 4\cos(\pi x)|\,\mathrm{d}x = \log 2 > 0,$$

this means that there is an $N \in \mathbb{N}$ such that $\frac{\log A_n}{n+1} > 0$ for $n \geq N$. Thus we have $A_n > 1$ for these n and this proves the theorem. q.e.d.

Remark II.4.14 We shall make a few remarks about Lemma II.4.12 and Theorem II.4.13.

- If the number $\xi = \frac{1}{\pi}\arccos\left(-\frac{1}{4}\right)$ is algebraic, Lioville's theorem (Theorem 0.5.20) immediately gives an upper bound for $\left|\xi - \frac{a}{b}\right|$ and we would not need Lemma II.4.12. Due to the Lindemann-Weierstraß theorem (Theorem 0.5.22), $\arccos\left(-\frac{1}{4}\right)$ is transcendental, thus ξ could be algebraic (although there is no reason whatsoever to believe that this is true).

- Lemma II.4.12 and the proof of Theorem II.4.13 show that ξ is a counterexample to the second statement of Lévy's theorem (Theorem I.3.4). When looking at the continued fraction expansion of ξ, the first 26 partial quotients are bounded by 5. This is evidence (although not a proof) for the fact that ξ cannot be too well approximated. In fact, among the first 1000 partial quotients there are exactly 35 partial quotients that are greater than 40, which is exactly the expected value due to Khintchine's theorem (Theorem I.3.3).

- We can show directly (i.e., without using the integrality of $\det(BL_n)$ and Λ_n) that the number V_n defined in Theorem II.4.13 is rational, using Galois theory: Note that

$$\begin{aligned}\lambda_{j,l} &= 1 + 2\cos\left(\frac{j\pi}{n+1}\right) + 2\cos\left(\frac{l\pi}{n+1}\right)\\ &= 1 + e^{\frac{j\pi}{n+1}} + e^{\frac{-j\pi}{n+1}} + e^{\frac{l\pi}{n+1}} + e^{\frac{-l\pi}{n+1}}\\ &= 1 + e^{\frac{2j\pi i}{2(n+1)}} + e^{\frac{-2j\pi i}{2(n+1)}} + e^{\frac{2l\pi i}{2(n+1)}} + e^{\frac{-2l\pi i}{2(n+1)}},\end{aligned}$$

i.e., $\lambda_{j,l} \in \mathbb{Q}(\zeta_{2(n+1)})$ where $\zeta_{2(n+1)}$ is a primitive $(2n+2)$-nd root of unity. The extension $\mathbb{Q}(\zeta_{2(n+1)})/\mathbb{Q}$ is a Galois extension and the $\mathbb{Q}$-automorphisms of $\mathbb{Q}(\zeta_{2(n+1)})$ are given by $\sigma_k(\zeta_{2(n+1)}) = \zeta_{2(n+1)}^k$ where $1 \leq k \leq 2n+2$, $\gcd(k, 2n+2) = 1$ (compare Example 0.5.37). Since the complex exponential function has period $2\pi i$, we get with Lemma 0.5.6

$$\begin{aligned}
\sigma_k(V_n) &= \prod_{j \neq l} \sigma_k(\lambda_{j,l}) \\
&= \prod_{j \neq l} \left(1 + e^{\frac{2jk\pi i}{2(n+1)}} + e^{\frac{-2jk\pi i}{2(n+1)}} + e^{\frac{2lk\pi i}{2(n+1)}} + e^{\frac{-2lk\pi i}{2(n+1)}}\right) \\
&= \prod_{j \neq l} \left(1 + e^{\frac{2j\pi i}{2(n+1)}} + e^{\frac{-2j\pi i}{2(n+1)}} + e^{\frac{2l\pi i}{2(n+1)}} + e^{\frac{-2l\pi i}{2(n+1)}}\right) \\
&= V_n
\end{aligned}$$

for all $\mathbb{Q}$-automorphisms σ_k of $\mathbb{Q}(\zeta_{2(n+1)})$. Thus V_n lies in the fixed field of $\mathrm{Gal}(\mathbb{Q}(\zeta_{2(n+1)})/\mathbb{Q})$ in $\mathbb{Q}(\zeta_{2(n+1)})$, i.e., $V_n \in \mathbb{Q}$ due to Theorem 0.5.38.

- It should be possible (although probably technical) to determine the number N in Theorem II.4.13 explicitely. To get a best possible value for N, one has to compare the rate of convergence of $R_L(n)$, $R_M(n)$, and $R_R(n)$ in the proof of Theorem II.4.13. To do this, one needs explicit bounds on the number δ in Lemma II.4.12. Some useful results can be found in [Lau08].

II.4.5 Solvability via prime(ideal) decomposition

The results of Section II.4.3 already yield a way to determine whether $BLO(n,k)$ is completely solvable:

1. Compute the determinant of BL_n.
2. Check if k is coprime to the determinant.

With the Euclidean algorithm it is easy to decide whether two integers are coprime, but in general it is hard to determine all integers that are coprime to a given integer m, since that requires prime factorization. It seems that in our case the prime decomposition of $\det(BL_n)$ is somehow "nice" in the sense that "small" prime factors occur with "greater" multiplicity (compare Tables II.4.1 and II.4.2). If this turns out to be true for all n (compare Section II.4.7), this would imply that, for given n, the integers k such that $BLO(n,k)$ is completely solvable could be determined relatively easily.

In this section we discuss another way of examining the complete solvability of $BLO(n,k)$. Regarding our last results, this new approach will be more of theoretical interest than of pratical use.

We will discuss a slightly more general setting. Fix $m \in \mathbb{N}$ and let $\zeta_m = e^{\frac{2\pi i}{m}}$, i.e., ζ_m is a primitive m-th root of unity. Let $\alpha := \zeta_m + \zeta_m^{-1} = 2\cos\left(\frac{2\pi}{m}\right)$. From the multiple angle formulae for cosine we know that $2\cos\left(\frac{2l\pi}{m}\right)$ is a polynomial in α for any $l \in \mathbb{N}$. Therefore the eigenvalues, and hence the determinant of BL_n, lie in the ring $\mathbb{Z}[\alpha]$ (which is the ring of integers of $\mathbb{Q}(\alpha)$, see [Was97, Proposition 2.16]). If $\mathbb{Z}[\alpha]$ is a unique factorization domain, this gives a first (not especially useful) algorithm to check whether $BLO(n,k)$ is completely solvable:

1. Compute the prime decomposition of k and of each eigenvalue $\lambda_{j,l}$ in $\mathbb{Z}[\alpha]$.
2. For each eigenvalue, check if the decomposition has (up to units) no prime in common with the decomposition of k.

As already mentioned, this is (in most cases) not suitable for practical purposes, since we would need to check whether $\mathbb{Z}[\alpha]$ is a unique factorization domain and we would need to know the units and primes in $\mathbb{Z}[\alpha]$. In general, it is hard to check if $\mathbb{Z}[\alpha]$ is a unique factorization domain. This is the case precisely when the class number of the field $\mathbb{Q}(\zeta_m + \zeta_m^{-1})$ is 1. But the class numbers of these fields are not known for primes $p \geq 71$, see [Sch03].

To avoid these obstructions, we will discuss another approach. Here we use Lemma II.4.5 and we deal only with $k = p$, where p is a prime.

Let R be the ring of integers of $\mathbb{Q}(\zeta_m)$ (then $\mathbb{Z}[\alpha] \subset R$). Suppose that p is inert in R (compare Theorem 0.5.32). Then p is also inert in $\mathbb{Z}[\alpha]$, which means that it cannot be the product of two or more eigenvalues. So either one of the eigenvalues is divisible (in $\mathbb{Z}$) by p, or the determinant is coprime to p. For the first case, we have the following lemma.

Lemma II.4.15 *Let $n \in \mathbb{N}$ and $\lambda_{j,l}$ be the eigenvalues of the matrix BL_n. Suppose that $\lambda_{j,l}$ is rational for some j,l. Then there are r,s with $\lambda_{r,s} = 0$ and thus* $\det(BL_n) = 0$.

Proof. If one of the eigenvalues is rational, Theorem II.4.6 says that $n+1$ is a multiple of 5 or 6 (this follows completely analogously to the proof of Theorem II.4.7). But then we already know that there is an eigenvalue that is 0. q.e.d.

Therefore, if the determinant is nonzero and p is inert, then p does not divide the determinant. We know (see [Was97, Theorem 2.13]) that p is inert in R if and

only if p is a primitive root modulo n. Together we have proved the following theorem.

Theorem II.4.16 *Let $n+1$ neither be a multiple of 5 nor a multiple of 6. If p is a primitive root modulo $2n+2$ for every prime p dividing k, then $BLO(n,k)$ is completely solvable.*

Of course, Theorem II.4.16 is not very useful in practice, since there are primitive roots modulo m if and only if $m \in \{1, 2, 4, p^k, 2p^k\}$ for some odd prime p and $k \in \mathbb{N}$.

It is noteworthy that it is not neccesary for p to be a primitive root modulo $m = 2n+2$ to get complete solvability for $BLO(n,p)$. This is (among other reasons, see below) due to the fact that p could be inert in $\mathbb{Z}[\alpha]$ but split in R. For example, let $n = p = 3$. Then $m = 8$ and there are no primitive roots modulo 8, hence no prime p is inert in R. Since $\zeta_8 = \frac{1+i}{\sqrt{2}}$, we get $\alpha = \sqrt{2}$. To determine the decomposition type of 3 in $\mathbb{Z}[\alpha]$, we can compute the Legendre symbol $\left(\frac{\Delta}{3}\right)$, where $\Delta = 8$ is the discriminant of the field extension $\mathbb{Q}(\alpha)/\mathbb{Q}$ (compare Theorems 0.5.26 und 0.5.33). Since $\left(\frac{8}{3}\right) = -1$, the prime 3 is inert in $\mathbb{Z}[\alpha]$. Thus, $BLO(3,3)$ is completely solvable despite the fact that 3 is not a primitive root modulo 8. We get the complete solvability of $BLO(3,3)$ also from $\det(BL_3) = -7$, cf. Table II.4.1. It would be nice to not only have a sufficient but also a neccesary condition for complete solvability, but this seems to get very technical.

If there is a prime p dividing k that is not inert, nothing can be said (at least not without more work). Possibly p can be a product of some of the eigenvalues. In the simplest case, p is the product of only two eigenvalues, i.e., $p = \lambda_{j,l}\lambda_{r,s}$. Expanding this product and using the product-to-sum identity for cosine, this gives (in general) sums of twelve cosines, more than can be handled with Theorem II.4.6.

II.4.6 Variants of Lights Out

As already mentioned, there are some variants of the original puzzle "Lights Out":

- Variations of the size of the board (for example, "Mini Lights Out" has a 4×4 board, "Lights Out Deluxe" has a 6×6 board). Of course, the board size can be varied even more.

- More colors (in "Lights Out 2000", each button has three states: red, green, or off).

- Different behaviour when pressing buttons. There are many ways to vary "Lights Out" by just changing how buttons vary their states. For example, the game "Mini Lights Out" identifies the left-hand side with the right-hand side of the board as well as the top with the bottom (i.e., considers the game to be without a border), so that every button has exactly four neighbours. We call this variant **unbounded**.

In the last sections, we have discussed "Lights Out" with extensions of board size and number of colors. In this section, we will briefly mention results about the unbounded variant. We will denote by $ULO(n,k)$ the "unbounded Lights Out" on an $n \times n$ board with k colors.

Example II.4.17 We consider the puzzle $ULO(3,5)$. On the board, we will denote the five colors by $0, 1, 2, 3$, and 4. We define the changing of colors as follows:

$$0 \to 1 \to 2 \to 3 \to 4 \to 0$$

When starting with the board

2	0	1
3	0	1
4	4	3

pressing the leftmost button in the top row yields

3	1	2
4	0	1
0	4	3

.

For obvious reasons we will always let $n \geq 3$ in the "unbounded" case. In general, we can proceed analogously to the puzzle $BLO(n,k)$.

Let

$$UL_n = \begin{pmatrix} K_n & I_n & 0 & \cdots & I_n \\ I_n & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & I_n \\ I_n & \cdots & 0 & I_n & K_n \end{pmatrix}$$

where

$$K_n := \begin{pmatrix} 1 & 1 & 0 & \cdots & 1 \\ 1 & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 1 & \cdots & 0 & 1 & 1 \end{pmatrix}, \text{ i.e., } (K_n)_{i,j} = \begin{cases} 1, & |i-j| \in \{0, 1, n-1\} \\ 0, & \text{otherwise} \end{cases}.$$

The matrix UL_n is the analogue of the matrix BL_n in the bounded case:

Proposition II.4.18 *Let n and k be natural numbers and let $n \geq 3$.*

1. *The game $ULO(n,k)$ is completely solvable if and only if $\det(UL_n)$ is invertible modulo k, i.e., if and only if $\gcd(\det(UL_n), k) = 1$.*
2. *Fix n. Then $ULO(n,k)$ is completely solvable for no k if and only if $\det(UL_n) = 0$, and $ULO(n,k)$ is completely solvable for all k if and only if $\det(UL_n) = \pm 1$.*

Again we determine the eigenvalues of UL_n to get the determinant of UL_n. The matrices K_n are **circulant** matrices, whose eigenvalues are (see [Gra])

$$\mu_j = 1 + e^{\frac{2\pi ji}{n}} + e^{\frac{2\pi ji(n-1)}{n}} = 1 + 2\cos\left(\frac{2j\pi}{n}\right).$$

Here i denotes the imaginary unit. Since $UL_n = K_n \otimes I_n + I_n \otimes K_n - I_{n^2}$, the eigenvalues of UL_n are

$$\mu_{j,l} = 1 + 2\left(\cos\left(\frac{2j\pi}{n}\right) + \cos\left(\frac{2l\pi}{n}\right)\right), \text{ where } 1 \leq j \leq n.$$

and we have

$$\det(UL_n) = \prod_{j=1}^{n}\prod_{l=1}^{n} \mu_{j,l}.$$

Tables II.4.3 and II.4.4 show the nonzero determinants of UL_n for $3 \leq n \leq 40$, respectively, the prime decomposition of its absolute value (computed with Mathematica).

n	$\det(BL_n)$	prime decomposition of $\lvert\det(BL_n)\rvert$
3	80	$2^4 \cdot 5^1$
4	-1215	$3^5 \cdot 5^1$
7	142805	$5^1 \cdot 13^4$
8	-7004233215	$3^5 \cdot 5^1 \cdot 7^8$
9	870763219280	$2^4 \cdot 5^1 \cdot 17^4 \cdot 19^4$
11	87789257318405	$5^1 \cdot 23^4 \cdot 89^4$
13	45421053339522658203125	$5^9 \cdot 53^4 \cdot 233^4$
14	$-705954787139532931683 4815$	$3^1 \cdot 5^1 \cdot 13^{16} \cdot 29^4$
16	$-199131097036890007894277961 94815$	$3^5 \cdot 5^1 \cdot 7^8 \cdot 17^{12} \cdot 47^4$
17	42950557989828549673287680	$2^{16} \cdot 5^1 \cdot 67^4 \cdot 1597^4$
19	940419421873808776346210289268985127605	$5^1 \cdot 37^{12} \cdot 113^4 \cdot 647^4$

Table II.4.3: The nonzero determinants $\det(UL_n)$ for $3 \leq n \leq 20$ and their prime decompositions.

n	prime decomposition of $\lvert\det(BL_n)\rvert$
21	$2^4 \cdot 5^1 \cdot 13^4 \cdot 41^8 \cdot 43^8 \cdot 379^4 \cdot 421^4$
22	$3^1 \cdot 5^1 \cdot 23^{20} \cdot 43^4 \cdot 89^4 \cdot 199^4$
23	$5^1 \cdot 47^{16} \cdot 137^8 \cdot 28657^4$
26	$3^{25} \cdot 5^{25} \cdot 53^{12} \cdot 79^4 \cdot 233^4 \cdot 521^4$
27	$2^4 \cdot 5^1 \cdot 17^4 \cdot 19^4 \cdot 53^{12} \cdot 107^8 \cdot 109^{12} \cdot 269^4$
28	$3^{29} \cdot 5^1 \cdot 13^{16} \cdot 29^{28} \cdot 251^4 \cdot 281^4$
29	$5^1 \cdot 59^{16} \cdot 173^8 \cdot 20011^4 \cdot 514229^4$
31	$2^{40} \cdot 5^1 \cdot 61^{16} \cdot 433^4 \cdot 557^4 \cdot 929^8 \cdot 2417^4$
32	$3^5 \cdot 5^1 \cdot 7^8 \cdot 17^{12} \cdot 31^{32} \cdot 47^4 \cdot 127^8 \cdot 223^4 \cdot 2207^4$
33	$2^{44} \cdot 5^1 \cdot 23^4 \cdot 67^{16} \cdot 89^4 \cdot 197^8 \cdot 397^8 \cdot 3499^4 \cdot 19801^4$
34	$2^{64} \cdot 3^1 \cdot 5^1 \cdot 67^{20} \cdot 101^8 \cdot 103^8 \cdot 409^4 \cdot 1597^4 \cdot 3571^4$
37	$5^1 \cdot 31^8 \cdot 73^{24} \cdot 149^{12} \cdot 443^8 \cdot 1553^8 \cdot 2221^4 \cdot 3331^4$
38	$3^1 \cdot 5^1 \cdot 37^{48} \cdot 113^{20} \cdot 191^8 \cdot 647^4 \cdot 9349^4$
39	$2^4 \cdot 5^9 \cdot 53^4 \cdot 79^{16} \cdot 157^8 \cdot 233^4 \cdot 311^8 \cdot 389^8 \cdot 467^8 \cdot 10531^4 \cdot 135721^4$

Table II.4.4: The prime decomposition of the nonzero determinants $\lvert\det(UL_n)\rvert$ for $21 \leq n \leq 40$.

We get the following results for UL_n and $ULO(n,k)$.

Theorem II.4.19 *Let n be a natural number with $n \geq 3$.*

- *The matrix UL_n is not invertible if and only if n is a multiple of 5 or 6.*
- *We have $\det(UL_n) \neq \pm 1$.*

Corollary II.4.20 *Let n and k be natural numbers with $n \geq 3$.*

- *If n is a multiple of 5 or 6 and k is arbitrary, then $ULO(n,k)$ is not completely solvable.*
- *For each n that is neither a multiple of 5 nor a multiple of 6, there is a k_1 such that $ULO(n,k_1)$ is completely solvable, and a k_2 such that $ULO(n,k_2)$ is not completely solvable.*
- *If n is neither a multiple of 5 nor a multiple of 6 and p is a primitive root modulo n for every prime p dividing k, then $ULO(n,k)$ is completely solvable.*

These results can be proved completely analogously to the results in the bounded case, with exception of the second part of Theorem II.4.19. This is an immediate consequence of the following lemma:

Lemma II.4.21 *[New78, Theorem 2]*
Let A be an $n \times n$ $(0,1)$-matrix with exactly k ones in each row and column. Then $\det(A)$ *is a multiple of $k \cdot \gcd(n,k)$.*

Finally, we also have the following result about specific "Lights Out" games.

Proposition II.4.22 *Consider any "Lights Out" game where the number of buttons whose illumination state is changed by pressing a button is the same for every button and coincides with the number of light states (for example $BLO(2,3)$ or $ULO(n,5)$ for any n). Then this game is not completely solvable.*

Proof. Denote by k the number of different colors and assign to each color the respective number in $\{0,\ldots,k-1\}$ (where again 0 is assigned to unlit buttons). We denote by Σ_A the sum of all numbers on the board A modulo k. Now suppose we start with a board A such that $\Sigma_A \neq 0$ (such a board exists, take, for example, a board where exactly one button is lit with color 1 and all other buttons are unlit). Then pressing any button will increase exactly k of the assigned numbers by exactly 1. Thus Σ_B of the new board B is $\Sigma_A + k \equiv \Sigma_A \not\equiv 0 \bmod k$. Hence Σ remains unchanged for any button pressed. But since Σ_0 of the all empty board is 0, the chosen starting board cannot be transformed into the all empty board. q.e.d.

This result also applies to certain variants of "Lights Out" not covered in this chapter. For example, consider an $n \times n$ board with n colors where each button changes the illumination state of all buttons in the respective row (and none other). Then this puzzle is not completely solvable due to Proposition II.4.22.

II.4.7 Future Work

There are various questions concerning "Lights Out" and the matrices involved that one could consider.

Tables II.4.1 to II.4.4 show some interesting patterns that one could examine. These concern the primes that occur in the prime decomposition of $\det(BL_n)$, respectively $\det(UL_n)$, as well as the respective exponents. The questions below originate from some striking values in Tables II.4.1 to II.4.4. It would be nice to know whether the observed patterns are random or whether there is a conceptual reason. To abbreviate, let L_n denote both matrices BL_n and UL_n.

- Let p^* be the largest prime divisor of $\det(L_n)$. Is there a small δ such that $p^* = \mathcal{O}(|\det(L_n)|^{\delta})$?
- Do twin prime pairs occur in the prime decomposition of $\det(L_n)$ more often than expected? If yes, is there a reason for this fact? Does every twin prime pair occur in the prime decomposition of $\det(L_n)$ for some n?
- Is there a reason why some exponents in the prime decomposition are relatively large? Can anything be said about the maximal size of the exponents that occur in the prime decomposition of $\det(L_n)$?
- The exponent of the largest prime divisor p^* of $\det(BL_n)$ is often 1, while the exponent of the biggest prime divisor p^* of $\det(UL_n)$ is often 4. Is there a reason for this?
- The exponents in the prime decomposition of $\det(UL_n)$ are often even or actually powers of 2. Is there a reason for this?

It would also be nice to have a characterization of the starting boards that are not solvable (analogous to the characterization given in [AF98] for the puzzle $BLO(5,2)$).

- Characterize the starting boards of $BLO(n,k)$ (or $ULO(n,k)$ or other variants) that are not solvable.

In Section II.4.5 we have seen a sufficient condition for complete solvability of $BLO(n,p)$ and we have seen that this condition is not necessary. It would be nice to have a condition that is both sufficient and necessary.

- Using decomposition over cyclotomic fields, find a necessary and sufficient condition for complete solvability of $BLO(n, p)$.

In Section II.4.6 we considered a variant of "Lights Out", i.e., a different behaviour when pressing buttons. There are much more variants of this kind.

- Examine the complete solvability for other variants of "Lights Out".

More of theoretical interest than of practical use would be an explicit number for which Theorem II.4.13 holds.

- Determine the number N in Theorem II.4.13 (compare Remark II.4.14).

Appendix

A.1

The Number 12

Why are there exactly twelve other areas for which we have shown connections to number theory? What is so special about 12? Is there even anything special about it?

There is a nice "proof" that shows that any natural number n is interesting: Let M be the set of all natural numbers that are not interesting and let n_0 be the minimal element in M. Then n_0 is the smallest natural number that is not interesting. Since this is an interesting property, n_0 itself is interesting, thus we have $M = \emptyset$.

Of course this relies heavily on when we call a natural number interesting. But for every natural number n there is some property of n that someone would find interesting. So what are interesting properties of 12?

Before stating three interesting properties, we exclude some trivialities. Since $12 = 2^2 \cdot 3$, 12 is the smallest natural number that can be written in the form p^2q for primes p, q. But this is not really interesting, since any natural number is the smallest natural number with some constructed property. We exclude such "constructed" properties.

The properties that we discuss are about sublime numbers, primes in arithmetic progressions and the Riemann ζ-function.

We start with sublime numbers. A natural number n is called **sublime** if both the number of its divisors $\tau(n)$ and its divisor sum $\sigma(n)$ are perfect numbers. Since $\tau(12) = 6$ and $\sigma(12) = 28$, the number 12 is sublime. Are there any more sublime numbers? Yes, but only one more is known (and it is not known whether there are more or even whether there are finitely many). We deduce some properties that sublime numbers need to have. The arguments below follow [Bro].

First suppose that both n and $\sigma(n)$ are even and let $v_2(n) = k$. We know (see Theorem 0.5.1) that $\sigma(n)$ is perfect if and only if $\sigma(n) = (2^s - 1)2^{s-1}$ such that $2^s - 1$ is prime. Since $\sigma(n) = \prod_{p|n} \frac{p^{v_p(n)+1}-1}{p-1}$ we have $2^{k+1} - 1 | \sigma(n)$, thus $2^{k+1} - 1$ has to be a prime and all the other factors of $\sigma(n)$ need to multiply to 2^k. Since

$$\sigma(p^v) = \sum_{j=0}^{v} p^j \equiv v + 1 \bmod 2,$$

this can only happen if all exponents of the odd primes are odd. Suppose that $p > 2$ and $v > 1$. In this case we have

$$\sigma(p^v) = (1 + p) \sum_{j=0}^{\frac{v-1}{2}} p^{2j}.$$

The sum is even if and only if $\frac{v-1}{2}$ is odd and in this case we get

$$\sigma(p^v) = (1 + p)(1 + p^2) \sum_{j=0}^{\frac{v-3}{4}} p^{4j}.$$

But if $1 + p = 2^r$ we have

$$1 + p^2 = 1 + (2^r - 1)^2 = 2 + 2^{2r} - 2^{r+1} = 2(2^{2r-1} - 2^r + 1)$$

and this is only a power of 2 if $r = p = 1$, which is a contradiction since p is a prime. Thus $p = 2^v - 1$ needs to be a Mersenne prime and all the exponents of the odd primes have to be 1. Note that we indeed have at least one Mersenne prime in this setting, since otherwise $\sigma(n)$ would be odd. If n has prime decomposition $2^k \prod_{j=1}^{l} p_j$, then $\tau(n) = (k + 1) \cdot 2^l$. This is perfect if and only if $k + 1 = 2^{l+1} - 1$ is a Mersenne prime.

Together, this gives the following theorem.

Theorem A.1.1 ([Bro]) *Let $n \in \mathbb{N}$. If n and $\sigma(n)$ are even, then n is sublime if and only if n has prime decomposition $n = 2^k \prod_{j=1}^{l} p_j$ where*

- $p_j = 2^{v_j} - 1$ *are Mersenne primes,*
- $\sum_{j=1}^{l} v_j = k$,
- $2^{k+1} - 1$ *is a Mersenne prime,*
- $2^{l+1} - 1$ *is a Mersenne prime,*
- $k + 1 = 2^{l+1} - 1$.

Looking up a list of known Mersenne primes (for example [MPS]), we find exactly four values of k such that both $k + 1$ and $2^{k+1} - 1$ are Mersenne primes, namely $2, 6, 30, 126$. But $6 = 2^3 - 2$ and $30 = 2^5 - 2$ cannot be written as the sum of $2(= 3 - 1)$, respectively $4(= 5 - 1)$, different exponents of Mersenne primes, thus these values do not give sublime numbers. For $k = 2$ we have $n = 2^2 \cdot 3 = 12$ and

- $3 = 2^2 - 1$ is a Mersenne primes,
- $\sum_{j=1}^{1} 2 = 2$,
- $2^3 - 1 = 7$ is a Mersenne prime,
- $2^2 - 1 = 3$ is a Mersenne prime,
- $2 + 1 = 2^{1+1} - 1$.

For $k = 126$ we have

$$\begin{aligned} n &= 2^{126} \cdot (2^{61} - 1)(2^{31} - 1)(2^{19} - 1)(2^7 - 1)(2^5 - 1)(2^3 - 1) \\ &= 6086555670238378989670371734243169622657830773351885970528324860512791691264 \end{aligned}$$

and

- $2^{61} - 1, 2^{31} - 1, 2^{19} - 1, 2^7 - 1, 2^5 - 1, 2^3 - 1$ are Mersenne primes,
- $61 + 31 + 19 + 7 + 5 + 3 = 126$,
- $2^{127} - 1$ is a Mersenne prime,
- $2^7 - 1$ is a Mersenne prime,
- $127 = 2^7 - 1$.

To find more even sublime numbers with even divisor sum one would need to check $2^{2147483647} - 1$ and higher Mersenne numbers for primality.

It is not clear if odd sublime numbers can exist. Suppose n is odd and $\sigma(n)$ is even and perfect. If $n = \prod_{j=1}^{l} p_j$ is squarefree, then $\tau(n) = 2^l$, which is not perfect. Thus there is a prime q with $v_q(n) \geq 2$, i.e., $n = q^r \prod_{j=1}^{l} p_j^{v_j}$ with $r \geq 2$. Then

$$\sigma(n) = \frac{q^{r+1} - 1}{q - 1} \prod_{j=1}^{l} \frac{p^{v_j+1} - 1}{p - 1}.$$

As shown above, the factor $\frac{q^{r+1}-1}{q-1}$ cannot be a power of 2 since $r \geq 2$. By the same arguments as above, n therefore needs to be of the form $n = q^r \prod_{j=1}^{l} p_j$. Now we can argue analogously to the case where n is even, but we further have the condition that $\frac{q^{r+1}-1}{q-1}$ is a Mersenne prime. Thus we conclude

Theorem A.1.2 ([Bro]) *Let $n \in \mathbb{N}$. If n is odd and $\sigma(n)$ is even, then n is sublime if and only if n has prime decomposition $n = q^r \prod_{j=1}^{l} p_j$ where*

- $p_j = 2^{v_j} - 1$ *are Mersenne primes,*
- $\frac{q^{r+1}-1}{q-1} = 2^{k+1} - 1$ *is a Mersenne prime,*
- $\sum_{j=1}^{l} v_j = k$,
- $2^{l+1} - 1$ *is a Mersenne prime,*
- $r + 1 = 2^{l+1} - 1$.

As already mentioned, no such n is known. If there are no odd perfect numbers, then all sublime numbers are given by the characterizations in Theorems A.1.1 and A.1.2. Hence 12 is one of only two known sublime numbers.

Let us turn our attention to primes in arithmetic progressions. From Dirichlet's theorem on primes in arithmetic progressions (Theorem 0.5.2) we know that there are infinitely many primes $p \equiv a \bmod m$ if $\gcd(a, m) = 1$. This can be proved with the use of Dirichlet L-functions, see [Brü95]. In some cases there are elementary proofs similar to Euclid's proof of the infinitude of the primes. One could ask in which cases such a proof exists.

To deal with this problem, we first need to have a precise definition of such a proof.

Definition A.1.3 Let $a, m \in \mathbb{N}$ with $\gcd(a, m) = 1$. A polynomial $f \in \mathbb{Z}[x]$ is called a **Euclidean polynomial** for a mod m if the following holds: For all but finitely many $n \in \mathbb{N}_0$ the prime factors p of $f(n)$ satisfy $p \equiv 1 \bmod m$ or $p \equiv a \bmod m$, where the latter case occurs infinitely often.

With such Euclidean polynomials **Euclidean proofs** of Dirichlet's theorem are possible. We want to know in which cases such a polynomial exists. There is a nice characterization:

Theorem A.1.4 *There is an Euclidean polynomial $f \in \mathbb{Z}[x]$ for a mod m if and only if $a^2 \equiv 1 \bmod m$.*

It was shown by Schur [Sch12] that the condition $a^2 \equiv 1 \bmod m$ is sufficient, the necessity was shown by M. R. Murty and Thain [MT06], cf. [Cona].

We are interested in those m such that an Euclidean proof exists for every a mod m with $\gcd(a, m) = 1$. In fact, these are only finitely many. Let m have prime decomposition $m = \prod p_j^{v_j}$ and suppose that $a^2 \equiv 1 \bmod m$. By the Chinese remainder theorem this is equivalent to $a^2 \equiv 1 \bmod p_j^{v_j}$ for all j. We show that no primes greater than 3 can occur in the prime decomposition (this is essentially due to the fact that in this case $2^2 \not\equiv 1 \bmod p$). Suppose that a prime $p \geq 5$ divides m. Choose an integer k such that $\gcd(kp + 2, m) = 1$ (interestingly we can use Dirichlet's theorem on primes in arithmetic progressions to guarantee the existence of such an integer k). Then $(2 + kp)^2 \equiv 4 \not\equiv 1 \bmod p$, thus we also have $(2 + kp)^2 \not\equiv 1 \bmod m$. Hence we have $m = 2^a 3^b$ for some $a, b \in \mathbb{N}_0$. If $m > 25$ we get $5^2 \not\equiv 1 \bmod m$. Thus the only remaining cases are $m \in \{1, 2, 3, 4, 6, 8, 9, 12, 16, 18, 24\}$. But $5^2 \equiv 7 \bmod 9$, $5^2 \equiv 9 \bmod 16$, and $5^2 \equiv 7 \bmod 18$, thus m cannot be $9, 16$, or 18. As one can easily check, all other values of m, i.e.,

$$m = 1, 2, 3, 4, 6, 8, 12, 24,$$

fulfill the condition $a^2 \equiv 1 \bmod m$ for all coprime a.

This is another interesting property of 12: It is the second biggest natural number m such that there exists an Euclidean proof for the infinitude of primes in the arithmetic progression $a + km$ for every a that is coprime to m. A complete Euclidean proof for the infinitude of primes in every arithmetic progression $a + 24k$ with $\gcd(a, 24) = 1$ (i.e., for the biggest natural number with this property) can be found in [BL65].

Now we discuss an interesting property of 12 related to the Riemann ζ-function. This is perhaps (at least of the three properties mentioned here) the most famous property of 12. In "fancy words", this property is: 12 is the additive inverse of the multiplicative inverse of $\zeta(-1)$. Or, to be more concise, $\zeta(-1) = -\frac{1}{12}$. How is this possible? Wouldn't this mean that $\sum_{n=1}^{\infty} n = -\frac{1}{12}$, i.e., the sum of all natural numbers equals $-\frac{1}{12}$? Surprisingly, there is indeed a notion of the value of infinite sums where this (almost) makes sense. Before considering this, recall that the Riemann ζ-function is defined as a sum only if $\Re(s) > 1$, which is clearly not true for $s = -1$. But we can use the functional equation in Theorem 0.5.18 for the Riemann ζ-function for $s = -1$ to get

$$\pi^{\frac{1}{2}}\Gamma\left(-\frac{1}{2}\right)\zeta(-1) = \pi^{-1}\Gamma(1)\zeta(2),$$

i.e.,

$$\zeta(-1) = \frac{\Gamma(1)}{\Gamma\left(-\frac{1}{2}\right)}\pi^{-\frac{3}{2}}\zeta(2).$$

Directly from the definition we get $\Gamma(1) = 1$. To compute $\Gamma\left(-\frac{1}{2}\right)$ we just state **Euler's reflection formula**, see [FB06]:

Theorem A.1.5 (Euler's reflection formula) *Let $s \in \mathbb{C}\backslash\mathbb{Z}$. Then*

$$\Gamma(1-s)\Gamma(s) = \frac{\pi}{\sin(\pi s)}.$$

Using this formula for $s = \frac{1}{2}$ we get $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$. Since Γ satisfies the functional equation $\Gamma(s+1) = s\Gamma(s)$ we have $\Gamma\left(-\frac{1}{2}\right) = -2\sqrt{\pi}$. The evaluation of $\zeta(2)$ is also known as the **Basel problem**. There are many different ways to show that $\zeta(2) = \frac{\pi^2}{6}$, see [Cha03] for some of them. Alltogether we get

$$\zeta(-1) = \frac{1}{-2\sqrt{\pi}}\pi^{-\frac{3}{2}}\frac{\pi^2}{6} = -\frac{1}{12}.$$

As already mentioned, this can also be obtained with a different summation method. There are many different summation methods, cf. [Kno22]. Some of them can assign values to infinite sums that diverge in the usual sense. There are two things worth noting. First, although some of these summation methods may work on a particular sum while others may not work, any applicable method gives the same value to a fixed infinite sum. Second, even with these summation

methods we cannot assign a value to $\sum_{n=1}^{\infty} n$ directly. But we can do this indirectly. We start with the definition of **Abel summability**, one of the most familiar summation methods.

Definition A.1.6 Let $\sum_{n=0}^{\infty} a_n$ be an infinite sum. If the power series $\sum_{n=0}^{\infty} a_n x^n$ has radius of convergence at least 1 and the left limit $\lim_{x \nearrow 1} \sum_{n=0}^{\infty} a_n x^n$ is $s \in \mathbb{R}$, then we call $\sum_{n=0}^{\infty} a_n$ **Abel summable** with value s.

The series $\sum_{n=0}^{\infty} n$ is not Abel summable, but $\sum_{n=0}^{\infty}(-1)^{n+1}n$ is, and we have

$$\begin{aligned}\sum_{n=0}^{\infty}(-1)^{n+1}nx^n &= x\sum_{n=0}^{\infty} n(-x)^{n-1} = -x\frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{n=0}^{\infty}(-x)^n\right)\\ &= -x\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{1+x}\right) = \frac{x}{(1+x)^2}\\ &\xrightarrow[x\nearrow 1]{} \frac{1}{4},\end{aligned}$$

thus the value of $\sum_{n=0}^{\infty}(-1)^{n+1}n$ according to Abel summability is $\frac{1}{4}$. If we now (just formally) subtract in the following way:

$$\begin{array}{rrcccccccc}
 & \sum_{n=0}^{\infty} n & = & 1 & +2 & +3 & +4 & +5 & +6 & + \cdots \\
- & 4\sum_{n=0}^{\infty} n & = & & -4 & & -8 & & -12 & - \cdots \\
\hline
= & -3\sum_{n=0}^{\infty} n & = & 1 & -2 & +3 & -4 & +5 & -6 & + \cdots
\end{array}$$

we find that $\sum_{n=0}^{\infty} n = -\frac{1}{3}\sum_{n=0}^{\infty}(-1)^{n+1}n = -\frac{1}{12}$. There are many more ways for "showing" this with formal manipulation of infinite series.

Apart from the classical summation methods, there is the notion of **Ramanujan summation**. With this method one can assign a value to $\sum_{n=1}^{\infty} n$ directly, cf. [Del02]:

$$\sum_{n=0}^{\infty} n = \int_0^0 f(t)\,\mathrm{d}t - \frac{1}{2}f(0) - \sum_{k=1}^{\infty}\frac{B_{2k}}{(2k)!}f^{(2k-1)}(0)$$

with $f(n) = n$. Here B_k are the Bernoulli numbers. Thus we get

$$\sum_{n=0}^{\infty} n = -\frac{B_2}{2} = -\frac{1}{12}.$$

We conclude our treatment with the remark that there are of course more interesting properties of the number 12. One very interesting property connects cusp forms, the abelianization of $\mathrm{SL}_2(\mathbb{Z})$ and lattice polygons. Since this would be to complicated to discuss here, we refer to [Wen] and [PRV00].

A.2

The Number of Solutions of a Linear Diophantine Equation with the Circle Method

In Chapter I.10 we discussed the circle method, with which we can get asymptotic formulae for the representation numbers of Diophantine equations. We show the first step of this method applied to the easiest Diophantine equation, namely a linear equation. A similar example can be found in [Els09]. Since this source is not publicly available, we will show how the circle method works in this case.

Consider the Diophantine equation

$$36x + 48y + 24z = 60. \tag{A.2.1}$$

We compute the number $R_5(60)$ of solutions (x, y, z) of Equation (A.2.1) such that $(x, y, z) \in [-5, 5]^3$. Using the circle method, we get the following: Let

$$f_1(\alpha) = \sum_{|k_1| \leq 5} e(36k_1\alpha), \qquad f_2(\alpha) = \sum_{|k_2| \leq 5} e(48k_2\alpha), \qquad f_3(\alpha) = \sum_{|k_3| \leq 5} e(24k_3\alpha)$$

with $e(x) = e^{2\pi i x}$. Then we have

$$\begin{aligned} R_5(60) &= \sum_{|k_1|\le 5}\sum_{|k_2|\le 5}\sum_{|k_3|\le 5}\int_0^1 e(\alpha(36k_1+48k_2+24k_3-60))\,\mathrm{d}\alpha \\ &= \int_0^1 f_1(\alpha)f_2(\alpha)f_3(\alpha)e(-\alpha n)\,\mathrm{d}\alpha. \end{aligned}$$

In this case, the integral can be evaluated relatively easily. Note that

$$\begin{aligned} \sin(m\pi\alpha)\sum_{|k|\le L} e(\alpha km) &= \sin(m\pi\alpha)\left(1+2\sum_{k=1}^{L}\cos(2m\pi k\alpha)\right) \\ &= \sin(m\pi\alpha)+2\sum_{k=1}^{L}\cos(2m\pi k\alpha)\sin(m\pi\alpha) \\ &= \sin(m\pi\alpha)+\sum_{k=1}^{L}(\sin((2k+1)m\pi\alpha)-\sin((2k-1)m\pi\alpha)) \\ &= \sin((2L+1)m\pi\alpha), \end{aligned}$$

i.e.,

$$\sum_{|k|\le L} e(\alpha km) = \frac{\sin((2L+1)m\pi\alpha)}{\sin(m\pi\alpha)}.$$

This is essentially **Lagrange's formula** (compare [GR15, Formula 1.342]). With this identity we get

$$R_5(60) = \int_0^1 \frac{\sin(396\pi\alpha)\sin(528\pi\alpha)\sin(264\pi\alpha)}{\sin(36\pi\alpha)\sin(48\pi\alpha)\sin(24\pi\alpha)} e(-60\alpha)\,\mathrm{d}\alpha.$$

Since $R_5(60)$ is real, we even get

$$R_5(60) = \int_0^1 \frac{\sin(396\pi\alpha)\sin(528\pi\alpha)\sin(264\pi\alpha)}{\sin(36\pi\alpha)\sin(48\pi\alpha)\sin(24\pi\alpha)} \cos(120\pi\alpha)\,\mathrm{d}\alpha.$$

With a computer algebra system this can be easily computed and we get $R_5(60) = 30$.

We can get the number of solutions also directly, without the use of the circle method. With standard methods, we find the general solution of the Diophantine equation $36x + 48y + 24z = 60$. This is given by

$$x = -5+20s+2t-8st, \quad y = 5-15s-2t+6st, \quad z = t, \text{ with } t \in \mathbb{Z}, s \in \frac{1}{5-2t}\mathbb{Z}.$$

If $(x,y,z) \in [-5,5]^3$, we get $t \in [-5,5]$ and $s \in [-1,\frac{3}{2}]$. Thus, we find the solutions

$$(1,3,-5),(5,0,-5),(-1,4,-4),(3,1,-4),(-3,5,-3),(1,2,-3),(5,-1,-3),$$
$$(-1,3,-2),(3,0,-2),(-3,4,-1),(1,1,-1),(5,-2,-1),(-5,5,0),(-1,2,0),$$
$$(3,-1,0),(-3,3,1),(1,0,1),(5,-3,1),(-5,4,2),(-1,1,2),(3,-2,2),(-3,2,3),$$
$$(1,-1,3),(5,-4,3),(-5,3,4),(-1,0,4),(3,-3,4),(-3,1,5),(1,-2,5),(5,-5,5)$$

and there are indeed 30 of them.

This shows how the integral formula in the circle method works. In the next step, one would have to determine suitable major and minor arcs. Since this is in general one of the hardest parts in the circle method (and we do not need it in our example), we will not treat the next steps of the circle method here.

A.3

Examples and Computations on Minimal Sets

In this appendix we will show some more results about minimal sets, in particular those of congruence classes. In Appendix A.3.1 some more minimal sets of congruence classes are shown. These have been computed with Algorithm II.1.1. In Appendix A.3.2 we take a look at the number of elements in these minimal sets with exactly k digits. These values will support Conjecture II.1.19. In Appendix A.3.3 we show some values of digit measures for some minimal sets.

A.3.1 Minimal Sets of Congruence Classes

The following list shows the minimal sets for all congruences classes for the modulus m for $2 \leq m \leq 8$. The minimal sets (for the congruence classes not treated in Chapter II.1) have been computed with Algorithm II.1.1.

$$\begin{aligned}
\mathcal{S}([0]_2) &= \{2,4,6,8,10,30,50,70,90\} \\
\mathcal{S}([1]_2) &= \{1,3,5,7,9\} \\
\mathcal{S}([0]_3) &= \{3,6,9,12,15,18,21,24,27,42,45,48,51,54,57,72,75,78,81,84,87, \\
&\qquad 111,114,117,141,144,147,171,174,177,222,225,228,252,255,258, \\
&\qquad 282,285,288,411,414,417,441,444,447,471,474,477,522,525,528,
\end{aligned}$$

$$552, 555, 558, 582, 585, 588, 711, 714, 717, 741, 744, 747, 771, 774, 777, 822, 825, 828, 852, 855, 858, 882, 885, 888\}$$

$$\mathcal{S}([1]_3) = \{1, 4, 7, 22, 25, 28, 52, 55, 58, 82, 85, 88\}$$

$$\mathcal{S}([2]_3) = \{2, 5, 8, 11, 14, 17, 41, 44, 47, 71, 74, 77\}$$

$$\mathcal{S}([0]_4) = \{4, 8, 12, 16, 20, 32, 36, 52, 56, 60, 72, 76, 92, 96, 100, 300, 500, 700, 900\}$$

$$\mathcal{S}([1]_4) = \{1, 5, 9, 33, 37, 73, 77\}$$

$$\mathcal{S}([2]_4) = \{2, 6, 10, 14, 18, 30, 34, 38, 50, 54, 58, 70, 74, 78, 90, 94, 98\}$$

$$\mathcal{S}([3]_4) = \{3, 7, 11, 15, 19, 51, 55, 59, 91, 95, 99\}$$

$$\mathcal{S}([0]_5) = \{5, 10, 20, 30, 40, 60, 70, 80, 90\}$$

$$\mathcal{S}([1]_5) = \{1, 6\}$$

$$\mathcal{S}([2]_5) = \{2, 7\}$$

$$\mathcal{S}([3]_5) = \{3, 8\}$$

$$\mathcal{S}([4]_5) = \{4, 9\}$$

$$\mathcal{S}([0]_6) = \{6, 12, 18, 24, 30, 42, 48, 54, 72, 78, 84, 90, 114, 144, 150, 174, 210, 222, 228, 252, 258, 270, 282, 288, 414, 444, 450, 474, 510, 522, 528, 552, 558, 570, 582, 588, 714, 744, 750, 774, 810, 822, 828, 852, 858, 870, 882, 888, 1110, 1170, 1410, 1470, 1710, 1770, 2250, 2550, 2850, 4110, 4170, 4410, 4470, 4710, 4770, 5250, 5550, 5850, 7110, 7170, 7410, 7470, 7710, 7770, 8250, 8550, 8850\},$$

$$\mathcal{S}([1]_6) = \{1, 7, 25, 43, 49, 55, 85, 223, 229, 283, 289, 445, 523, 529, 583, 589, 823, 829, 883, 889\}$$

$$\mathcal{S}([2]_6) = \{2, 8, 14, 44, 50, 56, 74, 110, 116, 170, 176, 410, 416, 470, 476, 554, 710, 716, 770, 776\}$$

$$\mathcal{S}([3]_6) = \{3, 9, 15, 21, 27, 45, 51, 57, 75, 81, 87, 111, 117, 141, 147, 171, 177, 225, 255, 285, 411, 417, 441, 447, 471, 477, 525, 555, 585, 711, 717, 741, 747, 771, 777, 825, 855, 885\}$$

$$\mathcal{S}([4]_6) = \{4, 10, 16, 22, 28, 52, 58, 70, 76, 82, 88, 112, 118, 172, 178, 250, 256, 550, 556, 712, 718, 772, 778, 850, 856\}$$

$$\mathcal{S}([5]_6) = \{5, 11, 17, 23, 29, 41, 47, 71, 77, 83, 89, 143, 149, 221, 227, 281, 287, 443, 449, 743, 749, 821, 827, 881, 887\}$$

$$\mathcal{S}([0]_7) = \{7, 14, 21, 28, 35, 42, 49, 56, 63, 84, 91, 98, 105, 112, 119, 126, 133, 161, 168, 182, 189, 196, 203, 224, 245, 252, 259, 266, 294, 301, 308, 322, 329,$$

336, 343, 364, 392, 399, 406, 413, 434, 441, 448, 455, 483, 504, 511, 518, 525, 532, 539, 553, 581, 588, 595, 602, 609, 616, 644, 651, 658, 665, 686, 805, 812, 819, 826, 833, 861, 868, 882, 889, 896, 903, 924, 945, 952, 959, 966, 994, 1001, 1008, 1022, 1029, 1036, 1092, 1099, 1106, 1113, 1155, 1183, 1225, 1232, 1239, 1253, 1295, 1302, 1309, 1316, 1386, 1652, 1659, 1666, 1806, 1813, 1855, 1883, 1925, 1932, 1939, 1953, 1995, 2002, 2009, 2044, 2065, 2205, 2226, 2233, 2296, 2436, 2443, 2464, 2534, 2555, 2604, 2625, 2695, 2905, 2926, 2933, 2996, 3003, 3024, 3066, 3094, 3206, 3234, 3304, 3612, 3619, 3626, 3661, 3668, 3682, 3689, 3696, 3906, 3934, 4004, 4011, 4018, 3311, 3318, 3332, 3339, 3381, 3388, 3416, 3444, 3486, 4053, 4081, 4088, 4116, 4151, 4158, 4165, 4186, 4333, 4361, 4368, 4403, 4445, 4466, 4501, 4508, 4543, 4816, 4851, 4858, 4865, 4886, 5005, 5012, 5019, 5033, 5082, 5089, 5103, 5152, 5159, 5222, 5229, 5243, 5292, 5299, 5313, 5334, 5341, 5348, 5383, 5502, 5509, 5544, 5551, 5558, 5803, 5852, 5859, 5922, 5929, 5943, 5992, 5999, 6006, 6041, 6048, 6055, 6111, 6118, 6125, 6181, 6188, 6195, 6405, 6461, 6468, 6524, 6545, 6552, 6559, 6594, 6601, 6608, 6622, 6629, 6664, 6692, 6699, 6811, 6818, 6825, 6881, 6888, 6895, 8001, 8008, 8022, 8029, 8036, 8092, 8099, 8106, 8113, 8155, 8183, 8225, 8232, 8239, 8253, 8295, 8302, 8309, 8316, 8386, 8652, 8659, 8666, 8806, 8813, 8855, 8883, 8925, 8932, 8939, 8953, 8995, 9002, 9009, 9044, 9065, 9205, 9226, 9233, 9296, 9436, 9443, 9464, 9534, 9555, 9604, 9625, 9695, 9905, 9926, 9933, 9996, 10003, 10066, 10311, 10318, 10381, 10388, 11011, 11018, 11081, 11088, 11116, 11151, 11158, 11165, 11186, 11501, 11508, 11816, 11851, 11858, 11865, 11886, 12222, 12229, 12292, 12299, 12922, 12929, 12992, 12999, 13006, 13111, 13118, 13181, 13188, 13811, 13818, 13881, 13888, 16555, 18011, 18018, 18081, 18088, 18116, 18151, 18158, 18165, 18186, 18501, 18508, 18816, 18851, 18858, 18865, 18886, 19222, 19229, 19292, 19299, 19922, 19929, 19992, 19999, 20006, 20055, 20622, 20629, 20692, 20699, 22022, 22029, 22092, 22099, 22225, 22232, 22239, 22253, 22295, 22302, 22309, 22925, 22932, 22939, 22953, 22995, 24444, 25333, 26005, 26222, 26229, 26292, 26299, 26922, 26929, 26992, 26999, 29022, 29029, 29092, 29099, 29225, 29232, 29239, 29253, 29295,

29302, 29309, 29925, 29932, 29939, 29953, 29995, 30002, 30009, 30044, 30233, 30933, 32004, 32333, 33033, 33103, 33313, 33334, 33341, 33348, 33383, 33803, 34111, 34118, 34181, 34188, 34811, 34818, 34881, 34888, 36666, 39004, 39333, 40005, 40033, 40544, 41111, 41118, 41181, 41188, 41811, 41818, 41881, 41888, 43666, 44044, 44436, 44443, 44464, 44604, 45003, 45444, 48111, 48118, 48181, 48188, 48811, 48818, 48881, 48888, 50001, 50008, 50022, 50029, 50092, 50099, 50155, 50855, 51002, 51009, 51555, 52444, 53333, 55055, 55405, 55524, 55545, 55552, 55559, 55594, 58002, 58009, 58555, 59444, 60004, 60011, 60018, 60081, 60088, 60466, 61222, 61229, 61292, 61299, 61922, 61929, 61992, 61999, 64001, 64008, 64666, 65555, 66066, 66206, 66612, 66619, 66626, 66661, 66668, 66682, 66689, 66696, 66906, 68222, 68229, 68292, 68299, 68922, 68929, 68992, 68999, 80003, 80066, 80311, 80318, 80381, 80388, 81011, 81018, 81081, 81088, 81116, 81151, 81158, 81165, 81186, 81501, 81508, 81816, 81851, 81858, 81865, 81886, 82222, 82229, 82292, 82299, 82922, 82929, 82992, 82999, 83006, 83111, 83118, 83181, 83188, 83811, 83818, 83881, 83888, 86555, 88011, 88018, 88081, 88088, 88116, 88151, 88158, 88165, 88186, 88501, 88508, 88816, 88851, 88858, 88865, 88886, 89222, 89229, 89292, 89299, 89922, 89929, 89992, 89999, 90006, 90055, 90622, 90629, 90692, 90699, 92022, 92029, 92092, 92099, 92225, 92232, 92239, 92253, 92295, 92302, 92309, 92925, 92932, 92939, 92953, 92995, 94444, 95333, 96005, 96222, 96229, 96292, 96299, 96922, 96929, 96992, 96999, 99022, 99029, 99092, 99099, 99225, 99232, 99239, 99253, 99295, 99302, 99309, 99925, 99932, 99939, 99953, 99995, 100002, 100009, 111111, 111118, 111181, 111188, 111811, 111818, 111881, 111888, 115003, 118111, 118118, 118181, 118188, 118811, 118818, 118881, 118888, 181111, 181118, 181181, 181188, 181811, 181818, 181881, 181888, 185003, 188111, 188118, 188181, 188188, 188811, 188818, 188881, 188888, 200004, 222222, 222229, 222292, 222299, 222922, 222929, 222992, 222999, 223006, 229222, 229229, 229292, 229299, 229922, 229929, 229992, 229999, 292222, 292229, 292292, 292299, 292922, 292929, 292992, 292999, 293006, 299222, 299229, 299292, 299299, 299922, 299929,

299992, 299999, 300006, 331002, 331009, 333333, 338002, 338009, 400001, 400008, 444444, 446005, 500003, 554001, 554008, 555555, 600005, 662004, 666666, 669004, 800002, 800009, 811111, 811118, 811181, 811188, 811811, 811818, 811881, 811888, 815003, 818111, 818118, 818181, 818188, 818811, 818818, 818881, 818888, 881111, 881118, 881181, 881188, 881811, 881818, 881881, 881888, 885003, 888111, 888118, 888181, 888188, 888811, 888818, 888881, 888888, 900004, 922222, 922229, 922292, 922299, 922922, 922929, 922992, 922999, 923006, 929222, 929229, 929292, 929299, 929922, 929929, 929992, 929999, 992222, 992229, 992292, 992299, 992922, 992929, 992992, 992999, 993006, 999222, 999229, 999292, 999299, 999922, 999929, 999992, 999999, 1000006, 2000005, 3000004, 4000003, 5000002, 5000009, 6000001, 6000008, 8000006, 9000005}

$\mathcal{S}([1]_7) = \{1, 8, 22, 29, 36, 43, 50, 57, 64, 92, 99, 204, 246, 253, 260, 267, 274, 302,$ 309, 323, 330, 337, 344, 372, 379, 393, 400, 407, 442, 449, 456, 470, 477, 526, 533, 554, 596, 603, 652, 659, 666, 673, 904, 946, 953, 960, 967, 974, 2003, 2066, 2073, 2444, 2556, 2633, 2703, 2766, 2773, 3004, 3053, 3074, 3200, 3207, 3270, 3277, 3333, 3354, 3452, 3459, 3704, 3753, 3774, 3900, 3907, 3970, 3977, 4026, 4054, 4096, 4404, 4446, 4460, 4467, 4474, 4544, 4726, 4754, 4796, 5244, 5342, 5349, 5552, 5559, 5566, 5944, 6000, 6007, 6056, 6070, 6077, 6553, 6602, 6609, 6623, 6630, 6637, 6672, 6679, 6693, 6700, 6707, 6756, 6770, 6777, 9003, 9066, 9073, 9444, 9556, 9633, 9703, 9766, 9773, 20000, 20007, 20056, 20070, 20077, 20700, 20707, 20756, 20770, 20777, 27000, 27007, 27056, 27070, 27077, 27700, 27707, 27756, 27770, 27777, 30003, 30073, 30703, 30773, 33342, 33349, 33552, 33559, 37003, 37073, 37703, 37773, 40244, 40552, 40559, 40944, 44066, 44444, 44766, 47244, 47552, 47559, 47944, 55553, 55623, 55693, 60026, 60096, 60726, 60796, 65556, 66200, 66207, 66270, 66277, 66333, 66900, 66907, 66970, 66977, 67026, 67096, 67726, 67796, 90000, 90007, 90056, 90070, 90077, 90700, 90707, 90756, 90770, 90777, 97000, 97007, 97056, 97070, 97077, 97700, 97707, 97756, 97770, 97777, 300000, 300007, 300070, 300077, 300700, 300707, 300770, 300777, 307000, 307007, 307070,

307077, 307700, 307707, 307770, 307777, 335553, 370000, 370007, 370070, 370077, 370700, 370707, 370770, 370777, 377000, 377007, 377070, 377077, 377700, 377707, 377770, 377777, 555556},

$\mathcal{S}([2]_7) = \{2, 9, 16, 30, 37, 44, 51, 58, 65, 86, 100, 107, 114, 135, 170, 177, 184, 331,$
338, 345, 366, 401, 408, 415, 436, 450, 457, 471, 478, 485, 506, 534, 555, 576, 604, 611, 618, 646, 660, 667, 674, 681, 688, 800, 807, 814, 835, 870, 877, 884, 1031, 1038, 1045, 1101, 1108, 1115, 1150, 1157, 1171, 1178, 1185, 1311, 1318, 1381, 1388, 1731, 1738, 1745, 1801, 1808, 1815, 1850, 1857, 1871, 1878, 1885, 3334, 3355, 3411, 3418, 3481, 3488, 3614, 3684, 4006, 4055, 4076, 4111, 4118, 4181, 4188, 4335, 4566, 4706, 4755, 4776, 4811, 4818, 4881, 4888, 5000, 5007, 5035, 5070, 5077, 5336, 5504, 5546, 5560, 5567, 5574, 5700, 5707, 5735, 5770, 5777, 6001, 6008, 6036, 6071, 6078, 6134, 6400, 6407, 6470, 6477, 6631, 6638, 6666, 6701, 6708, 6736, 6771, 6778, 6834, 8031, 8038, 8045, 8101, 8108, 8115, 8150, 8157, 8171, 8178, 8185, 8311, 8318, 8381, 8388, 8731, 8738, 8745, 8801, 8808, 8815, 8850, 8857, 8871, 8878, 8885, 10334, 10411, 10418, 10481, 10488, 11055, 11111, 11118, 11181, 11188, 11755, 11811, 11818, 11881, 11888, 17334, 17411, 17418, 17481, 17488, 18055, 18111, 18118, 18181, 18188, 18755, 18811, 18818, 18881, 18888, 33336, 33546, 40000, 40007, 40035, 40070, 40077, 40700, 40707, 40735, 40770, 40777, 47000, 47007, 47035, 47070, 47077, 47700, 47707, 47735, 47770, 47777, 50045, 50745, 53335, 55400, 55407, 55470, 55477, 55666, 57045, 57745, 60006, 60076, 60706, 60776, 66334, 66614, 66684, 67006, 67076, 67706, 67776, 80334, 80411, 80418, 80481, 80488, 81055, 81111, 81118, 81181, 81188, 81755, 81811, 81818, 81881, 81888, 87334, 87411, 87418, 87481, 87488, 88055, 88111, 88118, 88181, 88188, 88755, 88811, 88818, 88881, 88888, 333335, 600000, 600007, 600070, 600077, 600700, 600707, 600770, 600777, 607000, 607007, 607070, 607077, 607700, 607707, 607770, 607777, 663336, 670000, 670007, 670070, 670077, 670700, 670707, 670770, 670777, 677000, 677007, 677070, 677077, 677700, 677707, 677770, 677777}

$\mathcal{S}([3]_7) = \{3, 10, 17, 24, 45, 52, 59, 66, 80, 87, 94, 115, 122, 129, 164, 185, 192, 199,$
206, 220, 227, 255, 262, 269, 276, 290, 297, 402, 409, 416, 444, 472, 479,

486, 500, 507, 514, 556, 570, 577, 584, 605, 612, 619, 640, 647, 654, 675, 682, 689, 815, 822, 829, 864, 885, 892, 899, 906, 920, 927, 955, 962, 969, 976, 990, 997, 1116, 1144, 1186, 1256, 1655, 1816, 1844, 1886, 1956, 2005, 2012, 2019, 2075, 2082, 2089, 2215, 2222, 2229, 2285, 2292, 2299, 2516, 2586, 2600, 2607, 2670, 2677, 2705, 2712, 2719, 2775, 2782, 2789, 2915, 2922, 2929, 2985, 2992, 2999, 4000, 4007, 4014, 4070, 4077, 4084, 4112, 4119, 4182, 4189, 4406, 4420, 4427, 4462, 4469, 4476, 4490, 4497, 4700, 4707, 4714, 4770, 4777, 4784, 4812, 4819, 4882, 4889, 5015, 5064, 5085, 5155, 5505, 5540, 5547, 5554, 5575, 5715, 5764, 5785, 5855, 6002, 6009, 6044, 6072, 6079, 6114, 6184, 6422, 6429, 6492, 6499, 6555, 6702, 6709, 6744, 6772, 6779, 6814, 6884, 8116, 8144, 8186, 8256, 8655, 8816, 8844, 8886, 8956, 9005, 9012, 9019, 9075, 9082, 9089, 9215, 9222, 9229, 9285, 9292, 9299, 9516, 9586, 9600, 9607, 9670, 9677, 9705, 9712, 9719, 9775, 9782, 9789, 9915, 9922, 9929, 9985, 9992, 9999, 11112, 11119, 11182, 11189, 11462, 11469, 11812, 11819, 11882, 11889, 18112, 18119, 18182, 18189, 18462, 18469, 18812, 18819, 18882, 18889, 20002, 20009, 20072, 20079, 20702, 20709, 20772, 20779, 22116, 22186, 22256, 22816, 22886, 22956, 27002, 27009, 27072, 27079, 27702, 27709, 27772, 27779, 29116, 29186, 29256, 29816, 29886, 29956, 40064, 40764, 41114, 41184, 41814, 41884, 44222, 44229, 44292, 44299, 44600, 44607, 44670, 44677, 44922, 44929, 44992, 44999, 47064, 47764, 48114, 48184, 48814, 48884, 50116, 50186, 50655, 50816, 50886, 55044, 55555, 55744, 57116, 57186, 57655, 57816, 57886, 60000, 60007, 60014, 60070, 60077, 60084, 60700, 60707, 60714, 60770, 60777, 60784, 67000, 67007, 67014, 67070, 67077, 67084, 67700, 67707, 67714, 67770, 67777, 67784, 81112, 81119, 81182, 81189, 81462, 81469, 81812, 81819, 81882, 81889, 88112, 88119, 88182, 88189, 88462, 88469, 88812, 88819, 88882, 88889, 90002, 90009, 90072, 90079, 90702, 90709, 90772, 90779, 92116, 92186, 92256, 92816, 92886, 92956, 97002, 97009, 97072, 97079, 97702, 97709, 97772, 97779, 99116, 99186, 99256, 99816, 99886, 99956, 111114, 111184, 111814, 111884, 118114, 118184, 118814, 118884, 181114, 181184, 181814, 181884, 188114, 188184, 188814, 188884, 200000, 200007, 200070, 200077,

200700, 200707, 200770, 200777, 207000, 207007, 207070, 207077,
207700, 207707, 207770, 207777, 221112, 221119, 221182, 221189,
221812, 221819, 221882, 221889, 228112, 228119, 228182, 228189,
228812, 228819, 228882, 228889, 270000, 270007, 270070, 270077,
270700, 270707, 270770, 270777, 277000, 277007, 277070, 277077,
277700, 277707, 277770, 277777, 291112, 291119, 291182, 291189,
291812, 291819, 291882, 291889, 298112, 298119, 298182, 298189,
298812, 298819, 298882, 298889, 811114, 811184, 811814, 811884,
818114, 818184, 818814, 818884, 881114, 881184, 881814, 881884,
888114, 888184, 888814, 888884, 900000, 900007, 900070, 900077,
900700, 900707, 900770, 900777, 907000, 907007, 907070, 907077,
907700, 907707, 907770, 907777, 921112, 921119, 921182, 921189,
921812, 921819, 921882, 921889, 928112, 928119, 928182, 928189,
928812, 928819, 928882, 928889, 970000, 970007, 970070, 970077,
970700, 970707, 970770, 970777, 977000, 977007, 977070, 977077,
977700, 977707, 977770, 977777, 991112, 991119, 991182, 991189,
991812, 991819, 991882, 991889, 998112, 998119, 998182, 998189,
998812, 998819, 998882, 998889},

$\mathcal{S}([4]_7) = \{4, 11, 18, 25, 32, 39, 53, 60, 67, 81, 88, 95, 102, 109, 123, 130, 137, 165,$
172, 179, 193, 200, 207, 221, 228, 263, 270, 277, 291, 298, 305, 333, 361,
368, 375, 501, 508, 515, 522, 529, 550, 557, 571, 578, 585, 592, 599, 613,
655, 662, 669, 683, 802, 809, 823, 830, 837, 865, 872, 879, 893, 900, 907,
921, 928, 963, 970, 977, 991, 998, 1005, 1033, 1075, 1222, 1229, 1292,
1299, 1355, 1663, 1705, 1733, 1775, 1922, 1929, 1992, 1999, 2013, 2062,
2069, 2083, 2202, 2209, 2223, 2230, 2237, 2272, 2279, 2293, 2622, 2629,
2692, 2699, 2713, 2762, 2769, 2783, 2902, 2909, 2923, 2930, 2937, 2972,
2979, 2993, 3000, 3007, 3063, 3070, 3077, 3301, 3308, 3315, 3350, 3357,
3371, 3378, 3385, 3665, 3700, 3707, 3763, 3770, 3777, 5002, 5009, 5065,
5072, 5079, 5100, 5107, 5170, 5177, 5261, 5268, 5555, 5562, 5569, 5702,
5709, 5765, 5772, 5779, 5800, 5807, 5870, 5877, 5961, 5968, 6122, 6129,
6192, 6199, 6521, 6528, 6591, 6598, 6633, 6661, 6668, 6822, 6829, 6892,
6899, 8005, 8033, 8075, 8222, 8229, 8292, 8299, 8355, 8663, 8705, 8733,

8775, 8922, 8929, 8992, 8999, 9013, 9062, 9069, 9083, 9202, 9209, 9223, 9230, 9237, 9272, 9279, 9293, 9622, 9629, 9692, 9699, 9713, 9762, 9769, 9783, 9902, 9909, 9923, 9930, 9937, 9972, 9979, 9993, 10000, 10007, 10063, 10070, 10077, 10700, 10707, 10763, 10770, 10777, 17000, 17007, 17063, 17070, 17077, 17700, 17707, 17763, 17770, 17777, 20122, 20129, 20192, 20199, 20661, 20668, 20822, 20829, 20892, 20899, 22033, 22222, 22229, 22292, 22299, 22733, 22922, 22929, 22992, 22999, 27122, 27129, 27192, 27199, 27661, 27668, 27822, 27829, 27892, 27899, 29033, 29222, 29229, 29292, 29299, 29733, 29922, 29929, 29992, 29999, 30013, 30083, 30713, 30783, 33100, 33107, 33170, 33177, 33555, 33800, 33807, 33870, 33877, 36663, 37013, 37083, 37713, 37783, 50005, 50075, 50705, 50775, 55521, 55528, 55591, 55598, 55661, 55668, 57005, 57075, 57705, 57775, 66315, 66385, 66665, 80000, 80007, 80063, 80070, 80077, 80700, 80707, 80763, 80770, 80777, 87000, 87007, 87063, 87070, 87077, 87700, 87707, 87763, 87770, 87777, 90122, 90129, 90192, 90199, 90661, 90668, 90822, 90829, 90892, 90899, 92033, 92222, 92229, 92292, 92299, 92733, 92922, 92929, 92992, 92999, 97122, 97129, 97192, 97199, 97661, 97668, 97822, 97829, 97892, 97899, 99033, 99222, 99229, 99292, 99299, 99733, 99922, 99929, 99992, 99999, 500000, 500007, 500070, 500077, 500700, 500707, 500770, 500777, 507000, 507007, 507070, 507077, 507700, 507707, 507770, 507777, 556665, 570000, 570007, 570070, 570077, 570700, 570707, 570770, 570777, 577000, 577007, 577070, 577077, 577700, 577707, 577770, 577777, 666663}

$\mathcal{S}([5]_7) = \{5, 12, 19, 26, 33, 40, 47, 61, 68, 82, 89, 96, 103, 110, 117, 131, 138, 166,$ 173, 180, 187, 201, 208, 222, 229, 243, 271, 278, 292, 299, 306, 320, 327, 341, 348, 362, 369, 376, 390, 397, 411, 418, 432, 439, 446, 481, 488, 600, 607, 642, 649, 663, 670, 677, 803, 810, 817, 831, 838, 866, 873, 880, 887, 901, 908, 922, 929, 943, 971, 978, 992, 999, 1006, 1041, 1048, 1076, 1111, 1118, 1146, 1181, 1188, 1300, 1307, 1370, 1377, 1643, 1706, 1741, 1748, 1776, 1811, 1818, 1846, 1881, 1888, 2000, 2007, 2042, 2049, 2070, 2077, 2203, 2210, 2217, 2231, 2238, 2273, 2280, 2287, 2441, 2448, 2700, 2707, 2742, 2749, 2770, 2777, 2903, 2910, 2917, 2931, 2938, 2973, 2980, 2987,

3001, 3008, 3022, 3029, 3071, 3078, 3092, 3099, 3211, 3218, 3281, 3288, 3442, 3449, 3666, 3701, 3708, 3722, 3729, 3771, 3778, 3792, 3799, 3911, 3918, 3981, 3988, 4163, 4366, 4422, 4429, 4443, 4492, 4499, 4863, 6032, 6039, 6046, 6466, 6606, 6620, 6627, 6662, 6669, 6676, 6690, 6697, 6732, 6739, 6746, 8006, 8041, 8048, 8076, 8111, 8118, 8146, 8181, 8188, 8300, 8307, 8370, 8377, 8643, 8706, 8741, 8748, 8776, 8811, 8818, 8846, 8881, 8888, 9000, 9007, 9042, 9049, 9070, 9077, 9203, 9210, 9217, 9231, 9238, 9273, 9280, 9287, 9441, 9448, 9700, 9707, 9742, 9749, 9770, 9777, 9903, 9910, 9917, 9931, 9938, 9973, 9980, 9987, 10001, 10008, 10071, 10078, 10701, 10708, 10771, 10778, 11163, 11443, 11863, 17001, 17008, 17071, 17078, 17701, 17708, 17771, 17778, 18163, 18443, 18863, 20032, 20039, 20732, 20739, 22111, 22118, 22181, 22188, 22300, 22307, 22370, 22377, 22811, 22818, 22881, 22888, 24442, 24449, 27032, 27039, 27732, 27739, 29111, 29118, 29181, 29188, 29300, 29307, 29370, 29377, 29811, 29818, 29881, 29888, 30000, 30007, 30042, 30049, 30070, 30077, 30700, 30707, 30742, 30749, 30770, 30777, 37000, 37007, 37042, 37049, 37070, 37077, 37700, 37707, 37742, 37749, 37770, 37777, 44231, 44238, 44441, 44448, 44931, 44938, 60366, 60443, 66022, 66029, 66092, 66099, 66666, 66722, 66729, 66792, 66799, 67366, 67443, 80001, 80008, 80071, 80078, 80701, 80708, 80771, 80778, 81163, 81443, 81863, 87001, 87008, 87071, 87078, 87701, 87708, 87771, 87778, 88163, 88443, 88863, 90032, 90039, 90732, 90739, 92111, 92118, 92181, 92188, 92300, 92307, 92370, 92377, 92811, 92818, 92881, 92888, 94442, 94449, 97032, 97039, 97732, 97739, 99111, 99118, 99181, 99188, 99300, 99307, 99370, 99377, 99811, 99818, 99881, 99888, 100000, 100007, 100070, 100077, 100700, 100707, 100770, 100777, 107000, 107007, 107070, 107077, 107700, 107707, 107770, 107777, 114441, 114448, 170000, 170007, 170070, 170077, 170700, 170707, 170770, 170777, 177000, 177007, 177070, 177077, 177700, 177707, 177770, 177777, 184441, 184448, 444442, 444449, 800000, 800007, 800070, 800077, 800700, 800707, 800770, 800777, 807000, 807007, 807070, 807077, 807700, 807707, 807770, 807777, 814441, 814448, 870000, 870007, 870070, 870077, 870700, 870707, 870770,

870777, 877000, 877007, 877070, 877077, 877700, 877707, 877770, 877777, 884441, 884448},

$\mathcal{S}([6]_7) = \{6, 13, 20, 27, 34, 41, 48, 55, 83, 90, 97, 104, 111, 118, 125, 174, 181, 188, 195, 223, 244, 251, 258, 293, 300, 307, 321, 328, 335, 370, 377, 391, 398, 405, 433, 440, 447, 454, 475, 503, 510, 517, 524, 531, 538, 573, 580, 587, 594, 804, 811, 818, 825, 874, 881, 888, 895, 923, 944, 951, 958, 993, 1000, 1007, 1021, 1028, 1070, 1077, 1091, 1098, 1105, 1140, 1147, 1154, 1175, 1224, 1294, 1700, 1707, 1721, 1728, 1770, 1777, 1791, 1798, 1805, 1840, 1847, 1854, 1875, 1924, 1994, 2211, 2218, 2225, 2281, 2288, 2295, 2435, 2533, 2911, 2918, 2925, 2981, 2988, 2995, 3023, 3051, 3058, 3093, 3233, 3303, 3310, 3317, 3331, 3338, 3373, 3380, 3387, 3723, 3751, 3758, 3793, 3933, 4003, 4024, 4073, 4094, 4325, 4395, 4423, 4444, 4493, 4500, 4507, 4570, 4577, 4703, 4724, 4773, 4794, 5004, 5011, 5018, 5074, 5081, 5088, 5144, 5221, 5228, 5291, 5298, 5333, 5704, 5711, 5718, 5774, 5781, 5788, 5844, 5921, 5928, 5991, 5998, 8000, 8007, 8021, 8028, 8070, 8077, 8091, 8098, 8105, 8140, 8147, 8154, 8175, 8224, 8294, 8700, 8707, 8721, 8728, 8770, 8777, 8791, 8798, 8805, 8840, 8847, 8854, 8875, 8924, 8994, 9211, 9218, 9225, 9281, 9288, 9295, 9435, 9533, 9911, 9918, 9925, 9981, 9988, 9995, 10051, 10058, 10751, 10758, 11444, 11500, 11507, 11570, 11577, 12221, 12228, 12291, 12298, 12921, 12928, 12991, 12998, 17051, 17058, 17751, 17758, 18444, 18500, 18507, 18570, 18577, 19221, 19228, 19291, 19298, 19921, 19928, 19991, 19998, 22154, 22224, 22294, 22854, 22924, 22994, 29154, 29224, 29294, 29854, 29924, 29994, 30225, 30295, 30533, 30925, 30995, 33011, 33018, 33081, 33088, 33333, 33711, 33718, 33781, 33788, 37225, 37295, 37533, 37925, 37995, 40004, 40074, 40704, 40774, 44225, 44295, 44435, 44925, 44995, 47004, 47074, 47704, 47774, 50000, 50007, 50021, 50028, 50070, 50077, 50091, 50098, 50700, 50707, 50721, 50728, 50770, 50777, 50791, 50798, 57000, 57007, 57021, 57028, 57070, 57077, 57091, 57098, 57700, 57707, 57721, 57728, 57770, 57777, 57791, 57798, 80051, 80058, 80751, 80758, 81444, 81500, 81507, 81570, 81577, 82221, 82228, 82291, 82298, 82921, 82928, 82991, 82998, 87051, 87058, 87751, 87758, 88444, 88500, 88507, 88570, 88577, 89221, 89228, 89291,$

89298, 89921, 89928, 89991, 89998, 92154, 92224, 92294, 92854, 92924, 92994, 99154, 99224, 99294, 99854, 99924, 99994, 222221, 222228, 222291, 222298, 222921, 222928, 222991, 222998, 229221, 229228, 229291, 229298, 229921, 229928, 229991, 229998, 292221, 292228, 292291, 292298, 292921, 292928, 292991, 292998, 299221, 299228, 299291, 299298, 299921, 299928, 299991, 299998, 400000, 400007, 400070, 400077, 400700, 400707, 400770, 400777, 407000, 407007, 407070, 407077, 407700, 407707, 407770, 407777, 442224, 442294, 442924, 442994, 449224, 449294, 449924, 449994, 470000, 470007, 470070, 470077, 470700, 470707, 470770, 470777, 477000, 477007, 477070, 477077, 477700, 477707, 477770, 477777, 922221, 922228, 922291, 922298, 922921, 922928, 922991, 922998, 929221, 929228, 929291, 929298, 929921, 929928, 929991, 929998, 992221, 992228, 992291, 992298, 992921, 992928, 992991, 992998, 999221, 999228, 999291, 999298, 999921, 999928, 999991, 999998}

$\mathcal{S}([0]_8) = \{8, 16, 24, 32, 40, 56, 64, 72, 96, 104, 112, 120, 144, 152, 192, 200, 304, 336, 344, 360, 376, 504, 512, 520, 544, 552, 592, 600, 704, 736, 744, 760, 776, 904, 912, 920, 944, 952, 992, 1000, 3000, 5000, 7000, 9000\}$

$\mathcal{S}([1]_8) = \{1, 9, 25, 33, 57, 65, 73, 305, 345, 377, 385, 505, 545, 553, 585, 705, 745, 777, 785\}$

$\mathcal{S}([2]_8) = \{2, 10, 18, 34, 50, 58, 66, 74, 90, 98, 114, 146, 154, 194, 306, 330, 338, 370, 378, 386, 514, 546, 554, 594, 706, 730, 738, 770, 778, 786, 914, 946, 954, 994\}$,

$\mathcal{S}([3]_8) = \{3, 11, 19, 27, 51, 59, 67, 75, 91, 99, 107, 147, 155, 187, 507, 547, 555, 587, 707, 747, 771, 779, 787, 907, 947, 955, 987\}$

$\mathcal{S}([4]_8) = \{4, 12, 20, 28, 36, 52, 60, 68, 76, 92, 100, 108, 116, 156, 180, 188, 196, 300, 308, 332, 372, 380, 388, 500, 508, 516, 556, 580, 588, 596, 700, 708, 732, 772, 780, 788, 900, 908, 916, 956, 980, 988, 996\}$

$\mathcal{S}([5]_8) = \{5, 13, 21, 29, 37, 61, 69, 77, 93, 101, 109, 117, 141, 149, 181, 189, 197, 301, 309, 333, 341, 349, 381, 389, 701, 709, 733, 741, 749, 781, 789, 901, 909, 917, 941, 949, 981, 989, 997\}$

$\mathcal{S}([6]_8) = \{6, 14, 22, 30, 38, 54, 70, 78, 94, 102, 110, 118, 150, 158, 182, 190, 198, 334,$

$$342, 374, 502, 510, 518, 550, 558, 582, 590, 598, 734, 742, 774, 902, 910, 918, 950, 958, 982, 990, 998\}$$

$$\mathcal{S}([7]_8) = \{7, 15, 23, 31, 39, 55, 63, 95, 103, 111, 119, 143, 183, 191, 199, 303, 335, 343, 383, 503, 511, 519, 543, 583, 591, 599, 903, 911, 919, 943, 983, 991, 999\}$$

A.3.2 The Number of Elements in Minimal Sets of Congruence Classes

Table A.3.1 shows the number of elements in $\mathcal{S}([a]_m)$ that have exactly k digits for $2 \leq m \leq 13$ and some specific a. These numbers are known for more sets but we omit the presentation here for lack of space. The values have been computed with Algorithm II.1.1.

set	1	2	3	4	5	6	7	8	9	10	11	Σ
$[0]_2$	4	5	–	–	–	–	–	–	–	–	–	9
$[1]_2$	5	–	–	–	–	–	–	–	–	–	–	5
$[0]_3$	3	18	54	–	–	–	–	–	–	–	–	75
$[1]_3$	3	9	–	–	–	–	–	–	–	–	–	12
$[2]_3$	3	9	–	–	–	–	–	–	–	–	–	12
$[0]_4$	2	12	5	–	–	–	–	–	–	–	–	19
$[1]_4$	3	4	–	–	–	–	–	–	–	–	–	7
$[2]_4$	2	15	–	–	–	–	–	–	–	–	–	17
$[3]_4$	2	9	–	–	–	–	–	–	–	–	–	11
$[0]_5$	1	8	–	–	–	–	–	–	–	–	–	9
$[1]_5$	2	–	–	–	–	–	–	–	–	–	–	2
$[2]_5$	2	–	–	–	–	–	–	–	–	–	–	2
$[3]_5$	2	–	–	–	–	–	–	–	–	–	–	2
$[4]_5$	2	–	–	–	–	–	–	–	–	–	–	2
$[0]_6$	1	11	36	27	–	–	–	–	–	–	–	75
$[1]_6$	2	5	13	–	–	–	–	–	–	–	–	20
$[2]_6$	2	5	13	–	–	–	–	–	–	–	–	20
$[3]_6$	2	9	27	–	–	–	–	–	–	–	–	38
$[4]_6$	1	10	14	–	–	–	–	–	–	–	–	25
$[5]_6$	1	10	14	–	–	–	–	–	–	–	–	25
$[0]_7$	1	11	68	218	340	160	10	–	–	–	–	808

set	1	2	3	4	5	6	7	8	9	10	11	Σ
$[1]_7$	2	9	37	74	84	34	–	–	–	–	–	240
$[2]_7$	2	8	40	111	103	34	–	–	–	–	–	298
$[3]_7$	1	10	55	138	157	160	–	–	–	–	–	521
$[4]_7$	1	11	58	147	155	34	–	–	–	–	–	406
$[5]_7$	1	11	60	156	155	74	–	–	–	–	–	457
$[6]_7$	1	10	151	146	156	104	–	–	–	–	–	568
$[0]_8$	1	8	30	5	–	–	–	–	–	–	–	44
$[1]_8$	2	5	12	–	–	–	–	–	–	–	–	19
$[2]_8$	1	9	24	–	–	–	–	–	–	–	–	34
$[3]_8$	1	9	17	–	–	–	–	–	–	–	–	27
$[4]_8$	1	9	33	–	–	–	–	–	–	–	–	43
$[5]_8$	1	8	30	–	–	–	–	–	–	–	–	39
$[6]_8$	1	8	30	–	–	–	–	–	–	–	–	39
$[7]_8$	1	7	25	–	–	–	–	–	–	–	–	33
$[0]_9$	1	8	56	288	690	336	168	48	6	–	–	1601
$[1]_9$	1	7	36	85	51	27	8	1	–	–	–	216
$[2]_9$	1	7	36	85	51	27	8	1	–	–	–	216
$[3]_9$	1	7	36	90	15	3	–	–	–	–	–	152
$[4]_9$	1	7	36	85	51	27	8	1	–	–	–	216
$[5]_9$	1	7	36	85	51	27	8	1	–	–	–	216
$[6]_9$	1	7	36	90	15	3	–	–	–	–	–	152
$[7]_9$	1	7	36	85	51	27	8	1	–	–	–	216
$[8]_9$	1	7	36	85	51	27	8	1	–	–	–	216
$[0]_{10}$	0	9	–	–	–	–	–	–	–	–	–	9
$[1]_{10}$	1	–	–	–	–	–	–	–	–	–	–	1
$[0]_{11}$	0	9	64	350	510	46	–	–	–	–	–	979
$[1]_{11}$	1	7	45	170	215	155	100	45	10	1	–	749
$[2]_{11}$	1	7	45	166	200	148	88	31	10	1	–	697
$[3]_{11}$	1	7	44	169	197	155	100	45	10	1	–	729
$[0]_{12}$	0	8	35	87	27	–	–	–	–	–	–	157
$[4]_{12}$	1	5	18	21	–	–	–	–	–	–	–	45
$[5]_{12}$	1	5	18	21	–	–	–	–	–	–	–	45
$[0]_{13}$	0	7	58	347	1269	1698	901	303	34	–	–	4617
$[10]_{13}$	0	7	58	320	935	892	355	93	14	2	–	2676

Table A.3.1: The number of elements in $\mathcal{S}([a]_m)$ that have exactly k digits for $2 \leq m \leq 13$

Table A.3.2 shows the maximal number of digits in minimal sets for some $m \leq 100$. The first value gives the maximal number of digits in $\mathcal{S}([a]_m)$ for $a \neq 0$, the second value those for $a = 0$.

modulus	maximal number	modulus	maximal number	modulus	maximal number
2	1/2	12	4/5	25	2/3
3	2/3	13	9/10	30	3/4
4	2/3	14	7/8	32	5/6
5	1/2	15	3/4	40	3/4
6	3/4	16	4/5	50	2/3
7	6/7	17	16/17	60	4/5
8	3/4	18	9/10	70	7/8
9	8/9	19	18/19	80	4/5
10	1/2	20	2/3	90	9/10
11	10/6	21	11/11	100	2/3

Table A.3.2: The maximal number of digits in $\mathcal{S}([a]_m)$ for some $m \leq 100$

A.3.3 Digit Measures of some Minimal Sets

Table A.3.3 shows the rounded values of some digit measures for some minimal sets.

Set M	$\mu_c(\mathcal{S}(M))$	$\mu_g(\mathcal{S}(M))$	$\mu_h(\mathcal{S}(M))$	$\mu_z(\mathcal{S}(M))$
$\mathbb{N}$	0,1	0,009	0,05	0,000477465
$\mathbb{P}$	0,050623559	0,004050506	0,024226514	0,000213392
$\mathbb{N}\backslash\mathbb{P}$	0,068555556	0,0042107	0,030194444	0,000217167
$\boxed{2}$	0,070555556	0,0060305	0,034583333	0,000319024
$[0]_2$	0,05	0,00405	0,024074074	0,000213386
$[1]_2$	0,055555556	0,005	0,027777778	0,000265258
$[0]_3$	0,059333333	0,0031854	0,024833333	0,000163471
$[1]_3$	0,043333333	0,00309	0,02	0,000161277
$[2]_3$	0,043333333	0,00309	0,02	0,000161277
$[0]_4$	0,036111111	0,0021205	0,015694444	0,000108939
$[1]_4$	0,037777778	0,00304	0,018148148	0,000160098
$[2]_4$	0,038888889	0,00215	0,016666667	0,00010964
$[3]_4$	0,032222222	0,00209	0,014444444	0,000108225

Set M	$\mu_c(\mathcal{S}(M))$	$\mu_g(\mathcal{S}(M))$	$\mu_h(\mathcal{S}(M))$	$\mu_z(\mathcal{S}(M))$
$[0]_5$	$0,02$	$0,00108$	$0,008518519$	$5,49379 \cdot 10^{-5}$
$[1]_5$	$0,022222222$	$0,002$	$0,011111111$	$0,000106103$
$[2]_5$	$0,022222222$	$0,002$	$0,011111111$	$0,000106103$
$[3]_5$	$0,022222222$	$0,002$	$0,011111111$	$0,000106103$
$[4]_5$	$0,022222222$	$0,002$	$0,011111111$	$0,000106103$
$[0]_6$	$0,027633333$	$0,001113627$	$0,01068963$	$5,56933 \cdot 10^{-5}$
$[1]_6$	$0,029222222$	$0,0020513$	$0,013324074$	$0,000107299$
$[2]_6$	$0,029222222$	$0,0020513$	$0,013324074$	$0,000107299$
$[3]_6$	$0,035222222$	$0,0020927$	$0,015194444$	$0,000108261$
$[4]_6$	$0,023777778$	$0,0011014$	$0,009648148$	$5,54281 \cdot 10^{-5}$
$[5]_6$	$0,023777778$	$0,0011014$	$0,009648148$	$5,54281 \cdot 10^{-5}$
$[0]_7$	$0,033706778$	$0,001117021$	$0,012068479$	$5,57373 \cdot 10^{-5}$
$[1]_7$	$0,037252667$	$0,002093775$	$0,015652762$	$0,000108275$
$[2]_7$	$0,036907111$	$0,002084112$	$0,015451466$	$0,000108044$
$[3]_7$	$0,030058889$	$0,00110564$	$0,011125317$	$5,54836 \cdot 10^{-5}$
$[4]_7$	$0,031587111$	$0,001115949$	$0,011596651$	$5,57235 \cdot 10^{-5}$
$[5]_7$	$0,031913778$	$0,001116158$	$0,011672841$	$5,57262 \cdot 10^{-5}$
$[6]_7$	$0,040807111$	$0,001115248$	$0,013808688$	$5,5611 \cdot 10^{-5}$
$[0]_8$	$0,023388889$	$0,001083005$	$0,009362963$	$5,49778 \cdot 10^{-5}$
$[1]_8$	$0,029111111$	$0,0020512$	$0,013296296$	$0,000107298$
$[2]_8$	$0,023777778$	$0,0010924$	$0,009555556$	$5,52055 \cdot 10^{-5}$
$[3]_8$	$0,023$	$0,0010917$	$0,009361111$	$5,51963 \cdot 10^{-5}$
$[4]_8$	$0,024777778$	$0,0010933$	$0,009805556$	$5,52175 \cdot 10^{-5}$
$[5]_8$	$0,023333333$	$0,001083$	$0,009351852$	$5,49777 \cdot 10^{-5}$
$[6]_8$	$0,023333333$	$0,001083$	$0,009351852$	$5,49777 \cdot 10^{-5}$
$[7]_8$	$0,021666667$	$0,0010725$	$0,008842593$	$5,47353 \cdot 10^{-5}$
$[0]_9$	$0,030228143$	$0,001085895$	$0,010847425$	$5,50147 \cdot 10^{-5}$
$[1]_9$	$0,02389309$	$0,001073686$	$0,009346921$	$5,47506 \cdot 10^{-5}$
$[2]_9$	$0,02389309$	$0,001073686$	$0,009346921$	$5,47506 \cdot 10^{-5}$
$[3]_9$	$0,023905889$	$0,00107369$	$0,009350974$	$5,47507 \cdot 10^{-5}$
$[4]_9$	$0,02389309$	$0,001073686$	$0,009346921$	$5,47506 \cdot 10^{-5}$
$[5]_9$	$0,02389309$	$0,001073686$	$0,009346921$	$5,47506 \cdot 10^{-5}$
$[6]_9$	$0,023905889$	$0,00107369$	$0,009350974$	$5,47507 \cdot 10^{-5}$
$[7]_9$	$0,02389309$	$0,001073686$	$0,009346921$	$5,47506 \cdot 10^{-5}$
$[8]_9$	$0,02389309$	$0,001073686$	$0,009346921$	$5,47506 \cdot 10^{-5}$

Table A.3.3: Values of some digit measures

A.4

Examples on the Distribution of the Number of Solutions of Linear Equations

In this appendix we show some examples of the distribution of (II.3.3) with respect to the modulus. We will see that the distribution will sometimes not only depend on the determinant, but also on the Smith normal form of the involved matrix A. We will also see the correct values for the error terms R and R_y in Theorem II.3.7 for these examples.

Let

$$
\begin{aligned}
A &:= \operatorname{diag}(1,1,1,1,1,1,1,1,1,1,6,6,6,90) \in \mathcal{M}_{14,14}(\mathbb{Z}),\\
B &:= \operatorname{diag}(1,1,1,1,1,1,1,1,1,1,1,3,12,540) \in \mathcal{M}_{14,14}(\mathbb{Z}),\\
C &:= \operatorname{diag}(1,1,1,1,1,1,1,1,1,1,1,1,1,19\,440) \in \mathcal{M}_{14,14}(\mathbb{Z}),\\
D &:= \operatorname{diag}(1,1,1,1,1,1,1,1,1,1,1,1,1,30\,030) \in \mathcal{M}_{14,14}(\mathbb{Z}),\\
E &:= \operatorname{diag}(2,2,2,2,2,2,2,2,2,2,2,2,2,2) \in \mathcal{M}_{14,14}(\mathbb{Z}),\\
F &:= \operatorname{diag}(1,1,1,1,1,1,1,4,4,4,4,4,4,4) \in \mathcal{M}_{14,14}(\mathbb{Z}),\\
G &:= \operatorname{diag}(1,1,1,1,1,1,1,1,1,1,2,4,8,16,16) \in \mathcal{M}_{14,14}(\mathbb{Z}),\\
H &:= \operatorname{diag}(1,1,1,1,1,1,1,1,1,1,1,1,1,1,16\,384) \in \mathcal{M}_{14,14}(\mathbb{Z})
\end{aligned}
$$

and $m = 100\,000$. Let Δ_M denote the determinant of M. Then we have

$$\begin{aligned}
\Delta_A &= 19\,440 = 3^5 \cdot 2^4 \cdot 5,\\
\Delta_B &= 19\,440 = 3^5 \cdot 2^4 \cdot 5,\\
\Delta_C &= 19\,440 = 3^5 \cdot 2^4 \cdot 5,\\
\Delta_D &= 30\,030 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13,\\
\Delta_E &= 16\,384 = 2^{14},\\
\Delta_F &= 16\,384 = 2^{14},\\
\Delta_G &= 16\,384 = 2^{14},\\
\Delta_H &= 16\,384 = 2^{14}.
\end{aligned}$$

Figures A.4.1 to A.4.8 and Tables A.4.1 to A.4.3 show the distribution of solutions for the above matrices. In the figures the points have coordinates $(y, \#_y(M, m))$. Points with $\#_y(M, m) = 0$ are omitted.

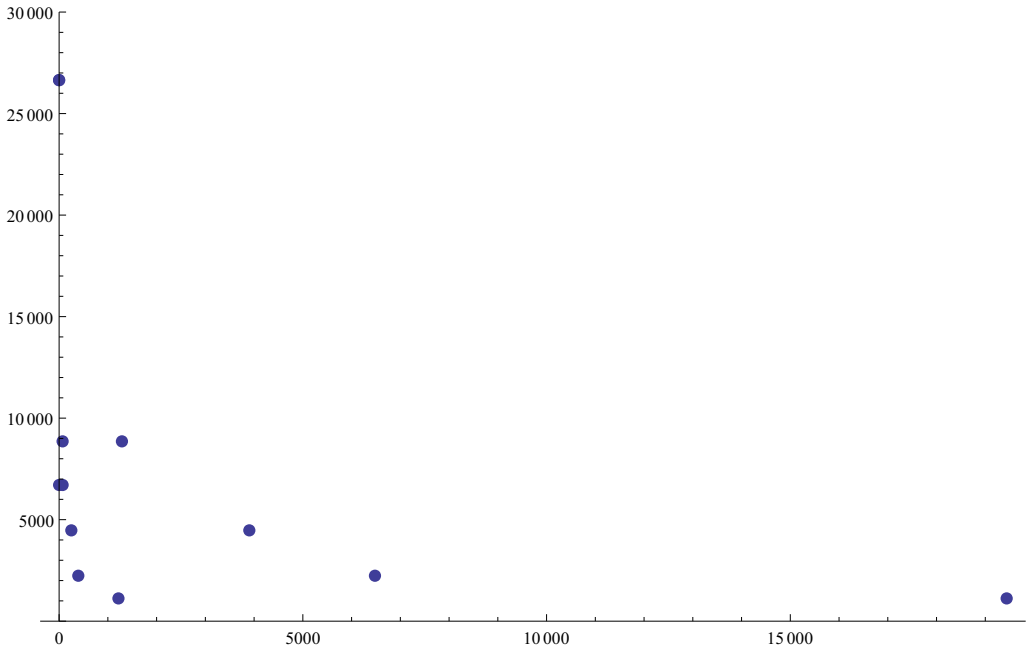

Figure A.4.1: The distribution of solutions for A.

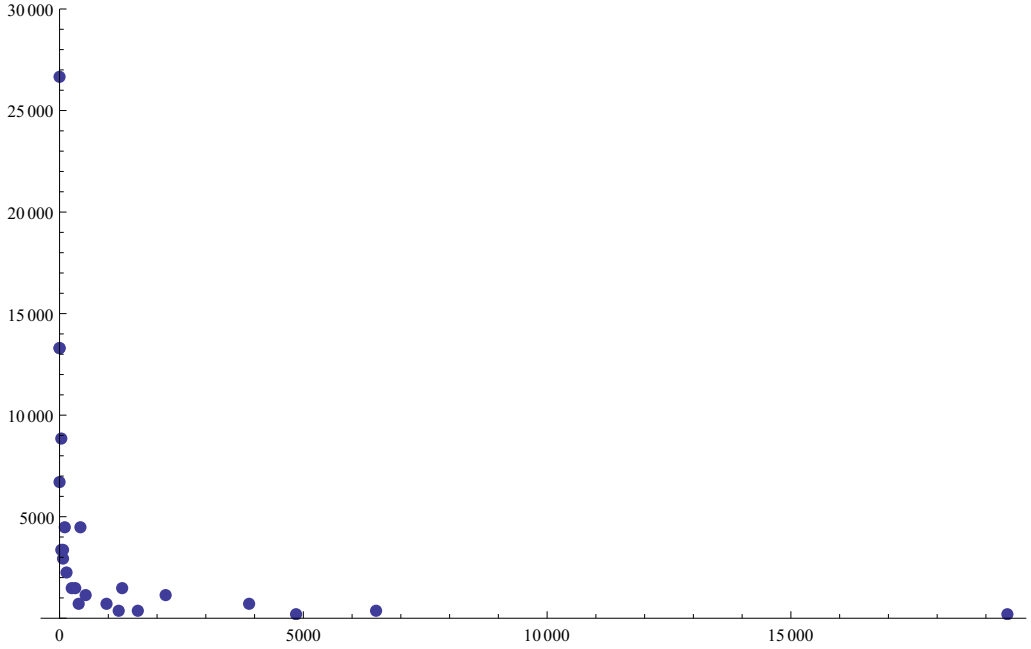

Figure A.4.2: The distribution of solutions for B.

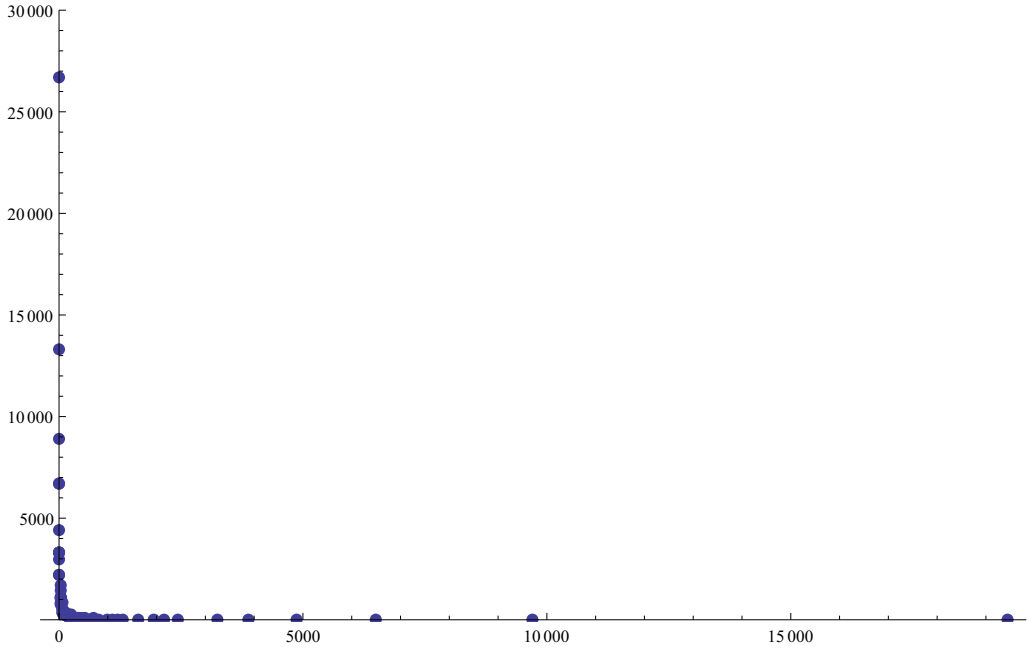

Figure A.4.3: The distribution of solutions for C.

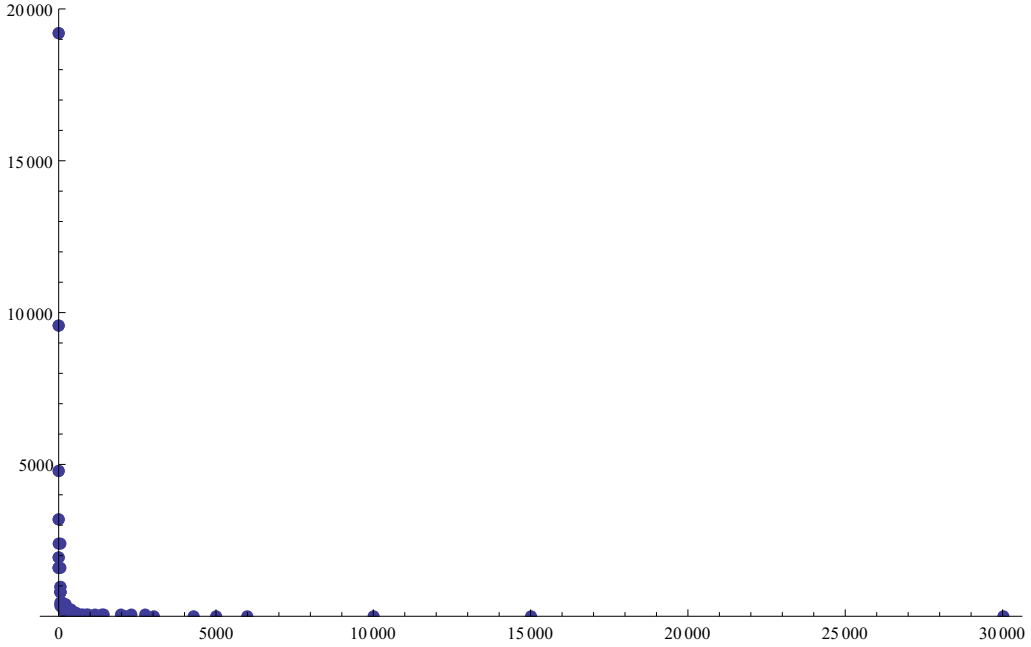

Figure A.4.4: The distribution of solutions for D.

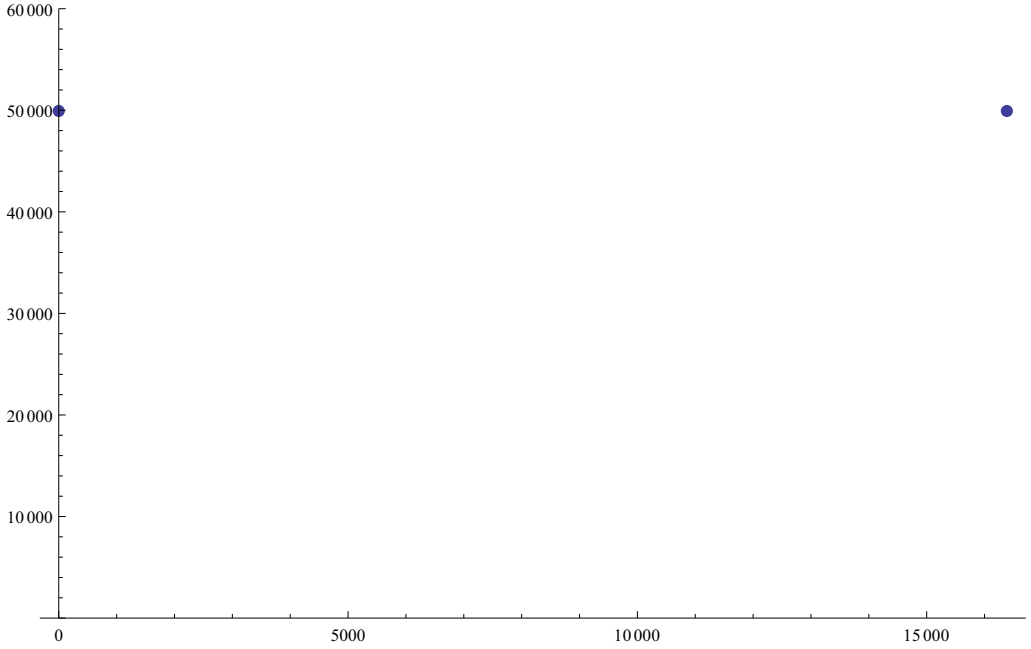

Figure A.4.5: The distribution of solutions for E.

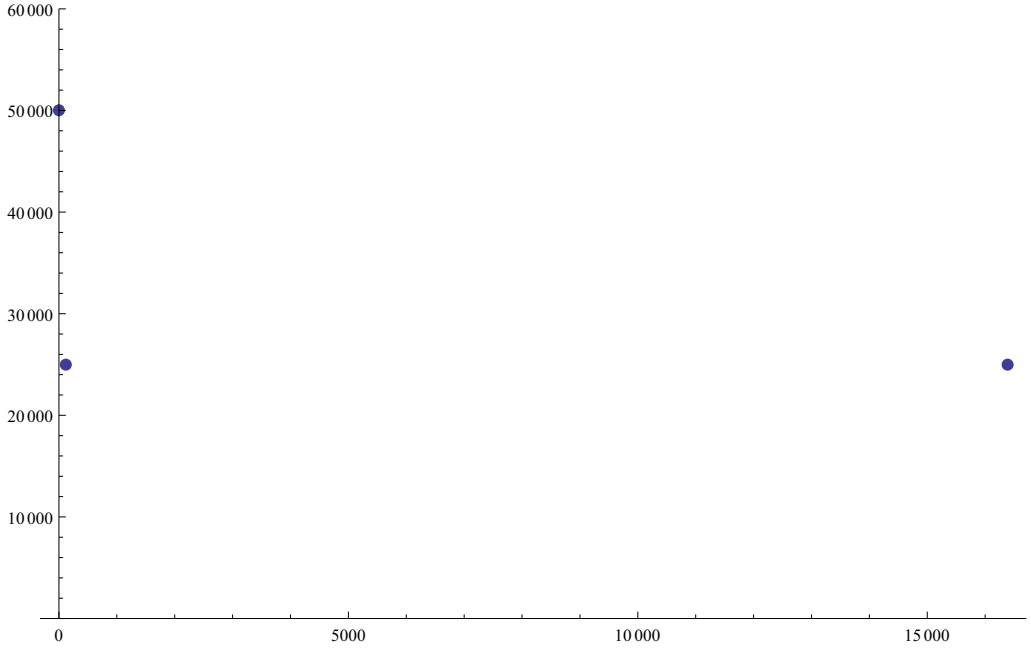

Figure A.4.6: The distribution of solutions for F.

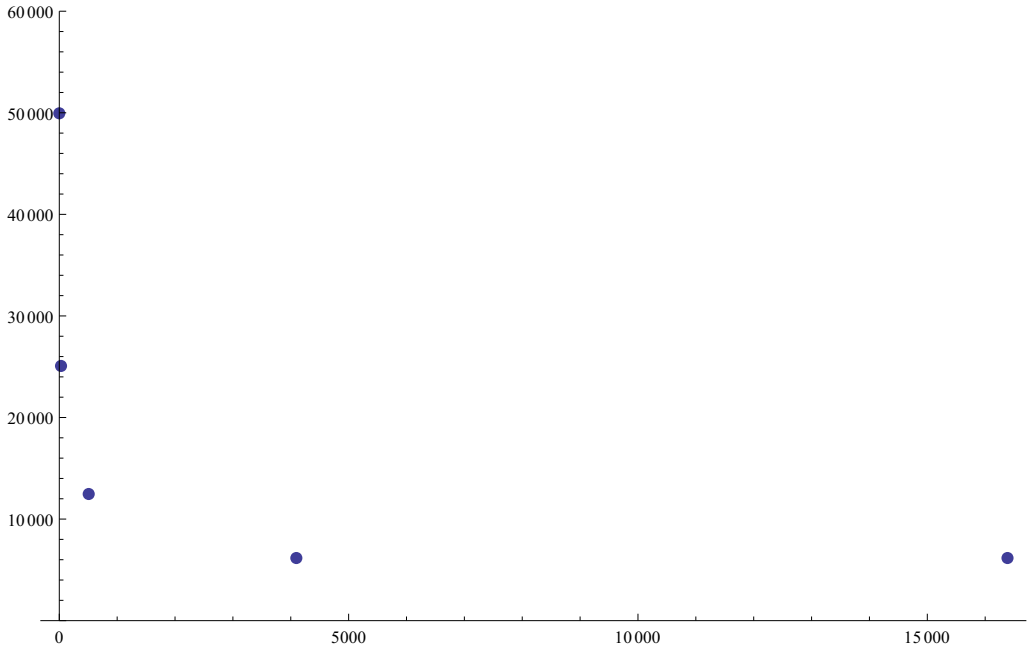

Figure A.4.7: The distribution of solutions for G.

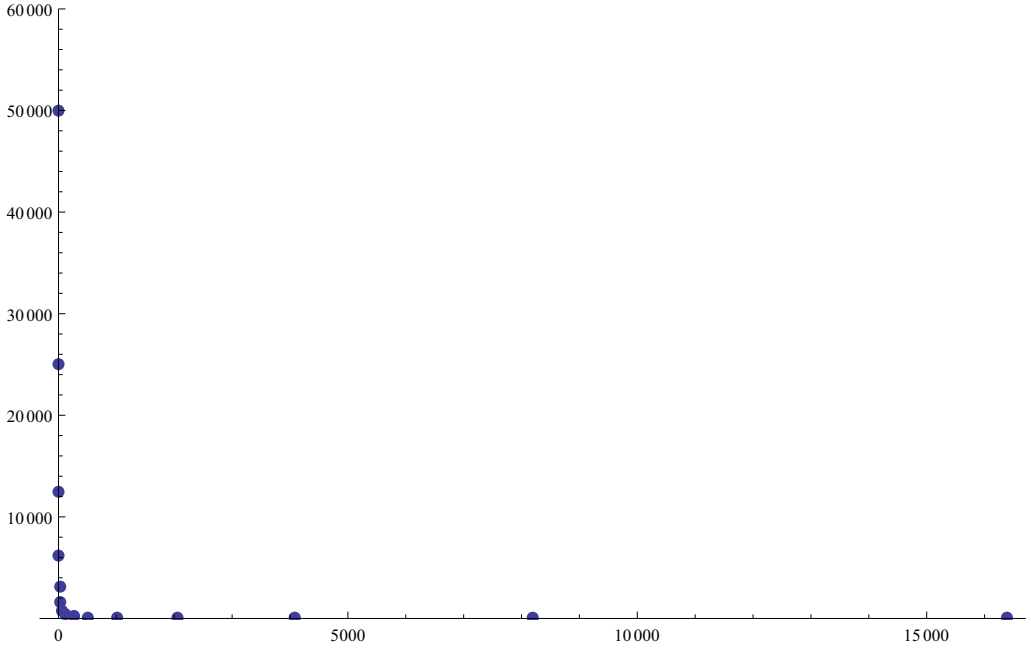

Figure A.4.8: The distribution of solutions for H.

Tables A.4.1 to A.4.3 show the exact values $\#_y(M,m)$ for some y and the matrices A to H. The values for R and R_y have been computed as the difference between these exact values and the asymptotic given in Theorem II.3.7.

The values for R and R_y are rounded to one decimal digit.

y	1	2	3	5	3^5	2^4	Δ
$\#_y(A,m)$	26 666	0	0	6 667	4 445	26 667	1 111
R/R_y	0.7	26 666.7	13 333.3	−0.3	−4 280.4	−23 333.7	−1 105.9
$\#_y(B,m)$	26 666	0	0	6 667	1 482	13 333	185
R/R_y	0.7	26 666.7	13 333.3	−0.3	−1 317.4	−9 999.7	−179.9
$\#_y(C,m)$	26 666	13 334	8 889	6 667	165	3 333	5
R/R_y	0.7	13 332.7	4 444.3	−0.3	−0.4	0.3	0.1

Table A.4.1: Values for the distribution of (II.3.3) for the matrices A to C.

y	1	2	3	5	7	11	13	Δ
$\#_y(D,m)$	19 181	19 182	9 590	4 795	3 196	1 918	1 598	3
R/R_y	−0.2	−1.2	0.4	0.2	0.8	0.1	0.4	0.3

Table A.4.2: Values for the distribution of (II.3.3) for the matrix D.

y	1	2	2^2	2^7	2^9	2^{10}	2^{12}	Δ
$\#_y(E,m)$	50 000	0	0	0	0	0	0	50 000
R/R_y	0	50 000	25 000	781.3	195.3	97.7	24.4	−49 993.9
$\#_y(F,m)$	50 000	0	0	25 000	0	0	0	25 000
R/R_y	0	50 000	25 000	−24 218.7	195.3	97.7	24.4	−24 993.9
$\#_y(G,m)$	50 000	0	0	0	12 500	0	6 250	6 250
R/R_y	0	50 000	25 000	781.3	−12 304.7	97.7	−6 225.6	−6 243.9
$\#_y(H,m)$	50 000	25 000	12 500	391	98	49	12	6
R/R_y	0	25 000	12 500	390.3	97.3	48.7	12.4	0.1

Table A.4.3: Values for the distribution of (II.3.3) for the matrices E to H.

Note that these tables include values y for which Theorem II.3.7 is not applicable. This amounts in large values of R and R_y. We also notice that the values of R and R_y in the cases where Theorem II.3.7 holds are small compared to the bound given in Theorem II.3.7. This means that better bounds could be possible (compare Section II.3.6).

List of Figures

List of Tables

List of Algorithms

List of Symbols

The list of symbols has the following sorting:

- The first block contains important sets of numbers such as the set of primes $\mathbb{P}$ or the upper half plane $\mathbb{H}$.
- The second block contains objects containing brackets (such as principal ideals) and absolute values. For example, the principal ideal (a) is sorted in this block, since the important symbol here is the bracket and not the "a". Tuples such as (M, τ) are not in this block, they are sorted according to the first letter inside the brackets.
- The third block contains special symbols (without letters) and integrals.
- The fourth block contains general operators and derived sets such as closures, sumsets, derivatives, and polynomial rings. For example, the sumset $A + B$ is not sorted alphabetically under "A" because the important part is the operator "$+$".
- The next blocks contain symbols sorted alphabetically. These are symbols that occur in Part 1 or in at least two chapters in Part 2. Greek letters are sorted at the end of the block corresponding to the first letter of their transliteration.
- The last four blocks contain operators and symbols used only a single chapter of Part 2.

If a symbol is defined in this thesis, the italic number at the end of the description indicates the page where the respective symbol is defined.

sets of numbers

$\mathbb{N}$

The set of natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$.

$\mathbb{N}_0$

The set of natural numbers with 0, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

$\mathbb{P}$

The set of primes $\mathbb{P} = \{2, 3, 5, 7, \ldots\}$.

$\mathbb{Z}$

The set of integers $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$.

$\mathbb{Q}$

The set of rational numbers.

$\mathbb{R}$

The set of real numbers.

$\mathbb{R}_E$

The set of real numbers that fulfills the condition E,
$\mathbb{R}_E = \{x \in \mathbb{R} : \text{the condition } E \text{ holds for } x\}$.

$\mathbb{C}$

The set of complex numbers.

$\mathbb{H}$

The upper half plane $\mathbb{H} = \{s \in \mathbb{C} : \Im(s) > 0\}$ *(p. 24)*.

$\mathbb{Z}/n\mathbb{Z}$

The quotient of the ring $\mathbb{Z}$ with respect to the ideal $n\mathbb{Z}$, i.e., the set of all residues modulo n (viewed as a group or a ring).

$\mathbb{F}_q$

The finite field with exactly q elements.

$\mathbb{Q}_p$

The p-adic numbers *(p. 63)*.

$\mathbb{Z}_p$

The p-adic integers *(p. 64)*.

brackets

(a)

The principal ideal generated by a.

$\left(\frac{a}{b}\right)$

The Kronecker symbol of a over b (equivalently, the Legendre symbol if b is prime) *(p. 39)*.

$\binom{V}{2}$

The set of all subsets of V that contain exactly two elements, $\binom{V}{2} = \{\{v, w\} : v, w, \in V\}$ *(p. 33)*.

$(x_0 : x_1 : x_2)$

Homogeneous coordinates in the projective plane *(p. 18)*.

$[L : K]$

The degree of the field extension L/K.

$\langle n \rangle_b$

The unique representation of n in base b *(p. 38)*.

$\langle a_0; a_1, a_2, \ldots \rangle$

The continued fraction with partial quotiens a_i *(p. 41)*.

$|A|$

The cardinality of the set A, i.e., its number of elements if A is finite.

$|\cdot|$ or $|\cdot|_\infty$

The usual absolute value.

$|\cdot|_p$

The p-adic absolute value *(p. 63)*.

$|\cdot|_T$

The absolute value on $\mathbb{F}_q(T)$ with $|T^e|_T = q^{-e}$ *(p. 67)*.

$\lfloor x \rfloor$

The floor function, $\lfloor x \rfloor = \max\{k \in \mathbb{Z} : k \leq x\}$ *(p. 42)*.

$\lceil x \rceil$

The ceiling function, $\lceil x \rceil = \min\{k \in \mathbb{Z} : k \geq x\}$ *(p. 90)*.

$\{x\}$

The fractional part of x, $\{x\} = x - \lfloor x \rfloor$ *(p. 42)*.

special symbols and integrals

$\infty_{\mathbb{H}}$

The point at infinity of $\mathbb{H}$ *(p. 27)*.

$\int_A f \, \mathrm{d}\mu$

The Lebesgue integral of the function f on the set A with respect to the measure μ.

general operators and derived sets

$\overline{A}$

The (topological) closure of A in a topological space (M, τ) *(p. 12)*.

$\overline{K}$

A fixed algebraic closure of the field K.

$\widehat{K}$

The completion of K with respect to a given valuation or absolute value *(p. 55)*.

K_v resp. $K_{|\cdot|}$

The completion of K with respect to the valuation v respectively the absolute value $|\cdot|$ *(p. 55)*.

$\widetilde{E}$

The reduction of the elliptic curve E modulo a given prime p *(p. 129)*.

$\overline{z}$

The complex conjugate of the complex number $z = x + iy$, i.e., $\overline{z} = x - iy$.

M^c

The complement of the set M in a given natural superset of M, mostly $\mathbb{R}$ or $\mathbb{N}$.

E^c

The complement of the edge set E of a graph $G = (V, E)$, $E^c = \binom{V}{2} \backslash E$ *(p. 34)*.

G^c

The complement of the graph G, $G^c = (V, E^c)$ *(p. 34)*.

K^G

The fixed field of G in K, i.e., $K^G := \{\alpha \in K : \sigma(\alpha) = \alpha \text{ for all } \sigma \in G\}$ *(p. 53)*.

$A + B$

The sumset of A and B, $A + B = \{a + b : a \in A, b \in B\}$ *(p. 42)*.

$A \oplus B$

The Kronecker sum of two matrices A and B *(p. 32)*.

$f \odot g$

The Hadamard product of two power series f and g *(p. 32)*.

$r \circ s$

The Hopf-Stiefel function $r \circ s = \beta_2(r, s)$ *(p. 90)*.

f'

The derivative of f. If f is a formal Laurent series, this is the formal derivative *(p. 32)*.

$f^{(n)}$

The n-th derivative of the function f.

$\frac{\mathrm{d}}{\mathrm{d}x} f(x)$

The derivative of a function $f : \mathbb{R} \to \mathbb{R}$.

$\frac{\partial}{\partial x_i} f(x_1, \ldots, x_n)$

The partial derivative of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ with respect to x_i.

$\dot{\gamma}(t)$

The derivative of a curve γ, i.e., $\dot{\gamma}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\gamma(t)$ *(p. 14)*.

$\mathfrak{a} \lhd R$

$\mathfrak{a}$ is an ideal of the ring R.

$f|_k M$

The transformed function $(f|_k M)(\tau) = \left(\frac{\mathrm{d}M\tau}{\mathrm{d}\tau}\right)^{\frac{k}{2}} f(M\tau)$ *(p. 107)*.

$R[x_1, \ldots, x_n]$

The polynomial ring in n variables over a ring R.

$R(x)$

The field of formal rational functions in one variable over a ring R.

$R[[x]]$

The ring of formal power series in one variable over a ring R *(p. 31)*.

$R((x))$

The field of formal Laurent series in one variable over a ring R *(p. 31)*.

R^*

The unit group of the ring R.

Σ^*

The set of all finite strings with symbols of Σ *(p. 58)*.

S^{-1}

The set of inverses of S (here S is a subset of some group G):
$S^{-1} = \{s^{-1} : s \in S\}$.

0 – 9

$\mathbf{0}_{|\cdot|}(K)$

The set of null sequences in the valued field $(K, |\cdot|)$ *(p. 55)*.

$\mathbf{0}_\infty(\mathbb{Q})$

The set of null sequences in $\mathbb{Q}$ with respect to the usual absolute value.

$\mathbf{0}_p(\mathbb{Q})$

The set of null sequences in $\mathbb{Q}$ with respect to the p-adic absolute value *(p. 64)*.

A

$A^2(K)$

The affine plane over some field K *(p. 18)*.

$a_f(m)$

The Fourier coefficients of the modular form f, $f(\tau) = \sum_{m\in\mathbb{Z}} a_f(m)e^{2\pi i m\tau}$ *(p. 108)*.

$\mathrm{Aut}(G)$

The group of automorphisms of the domain G (i.e., conformal functions $G \to G$) for $G \subset \widehat{\mathbb{C}}$ *(p. 23)*.

$\mathrm{Aut}(G)$

The group of automorphisms of the graph G *(p. 36)*.

B

$\mathscr{B}$

The Borel σ-algebra (of a given topological space) *(p. 56)*.

B_k

The k-th Bernoulli number defined via $\frac{t}{e^t-1} = \sum_{k=0}^{\infty} B_k \frac{t^k}{k!}$ *(p. 39)*.

$\beta_p(r,s)$

The function $\beta_p(r,s) = \min\{n : (x+y)^n = 0 \text{ in } \mathbb{F}_p[x,y]/(x^r,y^s)\}$ *(p. 90)*.

C

$\mathbf{c}_{|\cdot|}(K)$

The ring of Cauchy sequences in the valued field $(K, |\cdot|)$ *(p. 55)*.

$\mathbf{c}_\infty(\mathbb{Q})$

The ring of Cauchy sequences in $\mathbb{Q}$ with respect to the usual absolute value.

$\mathbf{c}_p(\mathbb{Q})$

The ring of Cauchy sequences in $\mathbb{Q}$ with respect to the p-adic absolute value *(p. 63)*.

$c(r,n)$

The Ramanujan sum $c(r,n) = \sum_{\substack{1\le k\le n\\(k,n)=1}} e^{\frac{2\pi irk}{n}}$ *(p. 84)*.

$C[a/b]$

The Ford circle with center at $\left(\frac{a}{b}, \frac{1}{2b^2}\right)$ and radius $\frac{1}{2b^2}$ *(p. 71)*.

$\widehat{\mathbb{C}}$

The Riemann sphere $\widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ *(p. 16)*.

$\mathbb{C}_p$

The completion of the algebraic closure of $\mathbb{Q}_p$ *(p. 67)*.

$\mathbb{C}_\infty$

The completion of the algebraic closure of the completion of $\mathbb{F}_q(t)$ *(p. 67)*.

$\mathrm{char}(K)$

The characteristic of the field K.

$\mathcal{C}l_K$

The ideal class group of $\mathcal{O}_K$, $\mathcal{C}l_K = I_K/P_K$ *(p. 51)*.

χ

A (Dirichlet or Hecke) character *(p. 82)*.

χ_A

The indicator function of the set A.

$\chi(D)$

The Euler charateristic of $D \subset M$ for a Riemannian manifold (M,g).

$\chi(G)$

The chromatic number of a graph G *(p. 35)*.

D

$d(v)$

The degree of a vertex v in a graph, $d(v) = |N(v)|$ *(p. 35)*.

$d(v, w)$

The distance between two vertices v, w in a graph,
$d(v, w) = \min\{l(P) : P \text{ is a path that connects } v \text{ and } w\}$ *(p. 35)*.

$d_{\mathbb{H}}$

The hyperbolic distance on the upper half plane $\mathbb{H}$ *(p. 26)*.

$\mathcal{D}_n$

The set of proper divisors of n, i.e., $\mathcal{D}_n = \{d : 1 \leq d \leq n-1, d|n\}$ *(p. 84)*.

$\deg(f)$

The degree (i.e., the biggest exponent) of the polynomial f.

$\det(A)$

The determinant of the matrix A.

$\operatorname{diam}(G)$

The diameter of a graph G, $\operatorname{diam}(G) = \max_{v,w \in V} d(v, w)\}$ *(p. 35)*.

$\triangle$

The triangle in $\mathbb{H}$ with ideal points $0, 1, \infty$ as corners *(p. 107)*.

Δ_E

The discriminant of the elliptic curve $y^2 = x^3 + ax + b$, $\Delta_E = 4a^3 + 27b^2$ *(p. 126)*.

Δ_f

The discriminant of the polynomial $f = x^n + a_{n-1} + \cdots + a_0$, i.e.,
$\Delta_f = \prod_{i<j}(\alpha_i - \alpha_j)^2 = (-1)^{\frac{n(n-1)}{2}} \prod_{i \neq j}(\alpha_i - \alpha_j)$ *(p. 31)*.

Δ_K

The discriminant of the number field K *(p. 48)*.

$\Delta(\tau)$

The modular discriminant.

∂D

The boundary of the set D in a topological space (M, τ).

$\partial \mathbb{H}$

The boundary at infinity of $\mathbb{H}$, $\partial \mathbb{H} = \mathbb{R} \cup \infty_{\mathbb{H}}$ *(p. 27)*.

E

$E(K)$

The points $(x, y) \in K^2$ that lie on the elliptic curve E *(p. 127)*.

$e(x)$

$e(x) = e^{2\pi i x}$ *(p. 19)*.

exp

The (real or p-adic) exponential function.

F

$\mathfrak{F}$

The fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}$ *(p. 107)*.

$\mathscr{F}_n$

The Farey sequence $\mathscr{F}_n$ *(p. 69)*.

$\widetilde{\mathscr{F}_n}$

The Farey sequence $\mathscr{F}_n$ without the fraction $\frac{0}{1}$ *(p. 70)*.

G

$G = (V, E)$

A graph with vertex set V and edge set E *(p. 33)*.

$G_n(d)$

The set $\{dk \in \mathbb{Z}/n\mathbb{Z} : k \in (\mathbb{Z}/n\mathbb{Z})^*\}$ *(p. 84)*.

$\mathrm{Gal}(K/\mathbb{Q})$

The Galois group of the Galois extension $K/\mathbb{Q}$ *(p. 53)*.

$\gcd(m, n)$

The greatest common divisor of m and n.

$\mathrm{GL}_n(R)$

The general linear group, i.e., $n \times n$ matrices over R (for a ring R) with nonzero determinant.

$\mathrm{GL}_2^+(\mathbb{R})$

The positive general linear group, i.e., 2×2 matrices over $\mathbb{R}$ with positive determinant *(p. 25)*.

Γ

The matrix group $\Gamma = \Gamma_0(1) = \mathrm{SL}_2(\mathbb{Z})$ *(p. 110)*.

$\Gamma_0(N)$

The subgroup of $\mathrm{SL}_2(\mathbb{Z})$ consisting of matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $c \equiv 0 \bmod N$ *(p. 110)*.

$\Gamma(s)$

The Γ-function $\Gamma(s) = \int_0^\infty e^{-t} t^{s-1} \, \mathrm{d}t$ *(p. 45)*.

H

h_K

The class number of the field K, $h_K = |Cl_K|$ *(p. 51)*.

I

$\Im(s)$

The imaginary part of the complex number s.

I_K

The group of all fractional ideals of $\mathcal{O}_K$ *(p. 51)*.

J

J_K

The idele group of K *(p. 82)*.

$J(n)$

The singular integral in the circle method *(p. 116)*.

$j(\tau)$

The j-invariant *(p. 109)*.

K

K

The curvature on a Riemannian manifold *(p. 14)*.

K

The Khintchine constant $K = \prod_{k=1}^{\infty} \left(1 + \frac{1}{k(k+2)}\right)^{\frac{\log k}{\log 2}} \approx 2.6854520010$ *(p. 76)*.

$K_m(C)$

The m-th K-group associated to the object C *(p. 80)*.

$\kappa(G)$

The vertex connectivity of the graph G *(p. 35)*.

$\kappa_g(s)$

The geodesic curvature at the point s on a Riemannian manifold (M, g).

L

$L(E, s)$

The Hasse-Weil L-function of the elliptic curve E *(p. 128)*.

$L(f, s)$

The L-function $L(f, s) = \sum_{m=1}^{\infty} a_f(m) m^{-s}$ attached to the modular form $f(\tau) = \sum_{m=0}^{\infty} a_f(m) e^{2\pi i m \tau}$ *(p. 110)*.

$L(s, \chi)$

The Dirichlet L-function attached to the Dirichlet character χ.

$L(s, \chi)$

The Hecke L-function attached to the Hecke character χ *(p. 82)*.

$L(M)$

The language accepted by the DFA M *(p. 59)*.

$L(\gamma)$

The length of a curve γ on a Riemannian manifold (M, g), i.e., $L(\gamma) = \int_a^b \sqrt{g(\gamma(t))(\dot{\gamma}(t), \dot{\gamma}(t))}\,\mathrm{d}t$ *(p. 14)*.

$l(P)$

The length of a path P in a graph, $l(P) = k$ if $P = v_0, \ldots, v_k$ *(p. 35)*.

$\mathcal{L}(M, \mathscr{A}, \mu)$

The set of integrable functions over the measure space $(M, \mathscr{A}, \mu)$ *(p. 57)*.

LCG_n

The coprime graph on $\{1, \ldots, n\}$ *(p. 86)*.

log

For real arguments the logarithmus naturalis, i.e., the logarithm with base e. For complex arguments the principal value of the complex logarithm with base e. For p-adic arguments the p-adic logarithm.

$\log_a$

The (real) logarithm with base a.

$\lambda(A)$

The Lebesgue measure of the set A.

M

$(M, \mathscr{A})$

A measurable space *(p. 55)*.

$(M, \mathscr{A}, \mu)$

A measure space *(p. 56)*.

$(M, \mathscr{A}, \mu, T)$

A dynamical system *(p. 56)*.

(M, g)

A Riemannian manifold, i.e., a differentiable manifold M together with a Riemannian metric g *(p. 14)*.

(M, τ)

A topological space, i.e., a set M equipped with a topology τ *(p. 11)*.

$M(n)$

Mertens function $M(n) = \sum_{k=1}^{n} \mu(k)$ *(p. 71)*.

$\mathcal{M}_{m,n}(R)$

The set of $m \times n$ matrices over R.

$\mathscr{M}_k$

The vector space of modular forms of weight k *(p. 109)*.

$\mathscr{M}_k(N)$

The vector space of modular forms of weight k and level N *(p. 110)*.

m_φ

The minimal polynomial of a linear map φ, i.e., the normalized integer polynomial f with smallest degree such that that $f(\varphi) = 0$.

$\mu(A)$

The measure of the set A *(p. 56)*.

$\mu(n)$

The Möbius function.

N

$N(v)$

The neighbourhood of a vertex v in a graph $G = (V, E)$:
$N(v) = \{w \in V : \{v, w\} \in E\}$ *(p. 35)*.

$N_k(v)$

The k-neighbourhood of a vertex v in a graph $G = (V, E)$:
$N_k(v) := \{w \in V : d(v, w) = k\}$ *(p. 35)*.

$N(\mathfrak{a})$

The norm of the ideal $\mathfrak{a}$ *(p. 52)*.

$N_K^L(x)$

The norm of $x \in L$ with respect to the field extension L/K *(p. 48)*.

O

$\mathcal{O}$

The point at infinity of an elliptic curve *(p. 126)*.

$\mathcal{O}_K$

The ring of integers of the number field K *(p. 47)*.

$\mathcal{O}_v$

The valuation ring of the completion K_v of a number field K, i.e., $\mathcal{O}_v = \{x \in K_v : v(x) \geq 0\}$ *(p. 81)*.

$\mathcal{O}(f)$

Landau $\mathcal{O}$, i.e., $g \in \mathcal{O}(f)$ if there is a $C > 0$ such that $|g(x)| \leq C\,|f(x)|$ for small enough or big enough x (depending on the context).

$\operatorname{ord}_{s_0}(f)$

The order of the meromorphic function f at s_0 *(p. 21)*.

$\omega(G)$

The clique number of a graph G *(p. 35)*.

P

P_K

The group of all principal ideals of $\mathcal{O}_K$ *(p. 51)*.

$P^2(K)$

The projective plane over some field K *(p. 18)*.

$\mathscr{P}(M)$

The power set of the set M.

Φ

The golden ratio $\Phi = \frac{1+\sqrt{5}}{2}$ *(p. 76)*.

$\varphi(n)$

The Euler totient function, $\varphi(n) = \sum_{(n,k)=1} 1$ *(p. 38)*.

π_q

Carlitz pi, $\pi_q = \prod_{k=1}^{\infty}\left(1 - \frac{T^{q^k}-T}{T^{q^{k+1}}-T}\right) \in \mathbb{C}_\infty$ *(p. 68)*.

Q

$(Q, \Sigma, \Delta, q_0, F)$

A DFA *(p. 58)*.

$(Q, \Sigma, \Delta, q_0, \Delta, \eta)$

A DFAO *(p. 59)*.

$\mathbb{Q}(\sqrt{d})$

The quadratic number field $\mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} : a, b \in \mathbb{Q}\}$ *(p. 49)*.

$\mathbb{Q}(\zeta_p)^{\mathbb{R}}$

The maximal totally real subfield of $\mathbb{Q}(\zeta_p)$ *(p. 81)*.

$\mathrm{Quot}(R)$

The quotient field of the ring R.

R

$\Re(s)$

The real part of the complex number s.

$R(n)$

The number of solutions of the Diophantine equation $D(x_1, \ldots, x_k) = n$, i.e., $R(n) = \lim_{N\to\infty} R_N(n)$ *(p. 115)*.

$R_N(n)$

The number of solutions of the Diophantine equation $D(x_1, \ldots, x_k) = n$ with $|x_i| \leq N$ for all i *(p. 115)*.

$R_s(n)$

The number of possible ways of writing n as a sum of s elements of a given set A *(p. 114)*.

$\mathrm{Res}(f; s_0)$

The residue of the meromorphic function f at s_0 *(p. 22)*.

S

$\mathscr{S}^2$

The sphere $\mathscr{S}^2 = \{s \in \mathbb{R}^3 : s_1^2 + s_2^2 + s_3^2 = 1\}$ *(p. 15)*.

$\mathscr{S}_k$

The vector space of cusp forms of weight k *(p. 109)*.

$\mathscr{S}_k(N)$

The vector space of cusp forms of weight k and level N *(p. 110)*.

$\mathfrak{S}(n)$

The singular series in the circle method *(p. 116)*.

$\mathrm{SL}_2(\mathbb{Z})$

The special linear group, i.e., 2×2 integer matrices with determinant 1.

$\sigma(n)$

The divisor sum $\sigma(n) = \sum_{d|n} d$ *(p. 38)*.

T

T_p

The Hecke operator $(T_p f)(\tau) = p^{k-1} f(p\tau) + \frac{1}{p} \sum_{b=1}^{p} f\left(\frac{\tau+b}{p}\right)$ *(p. 110)*.

$\tau(n)$

The number of divisors of n, $\tau(n) = \sum_{k|n} 1$ *(p. 38)*.

V

$V(f)$

The vanishing points of a polynomial f, i.e., all values c in affine plane or projective plane such that $f(c) = 0$ *(p. 18)*.

V_K

The adele ring of K *(p. 81)*.

$v_p(n)$

The exponent of p in the prime decomposition of $n \in \mathbb{N}$ *(p. 38)*.

$\mathrm{vol}(\Gamma)$

The volume of the lattice Γ. If $\Gamma = \mathbb{Z}v_1 + \cdots + \mathbb{Z}v_m$, then we have $\mathrm{vol}(\Gamma) = |\det(v_1, \ldots, v_n)|$ *(p. 29)*.

X

$X(H, S)$

The Cayley graph for the group H and its subset S *(p. 35)*.

X_n

The unitary Cayley graph $X_n = X(\mathbb{Z}/n\mathbb{Z}, (\mathbb{Z}/n\mathbb{Z})^*)$ *(p. 83)*.

$X_n(D)$

The gcd graph associated to the set $D \subset \mathcal{D}_n$ *(p. 84)*.

x_N resp. x_S

The stereographic projections $x_N(s_1, s_2, s_3) = \frac{1}{1+s_3}(s_1, s_2)$ respectively $x_S(s_1, s_2, s_3) = \frac{1}{1-s_3}(s_1, s_2)$ *(p. 15)*.

Z

$\mathbb{Z}[\sqrt{d}]$

The ring $\mathbb{Z}[\sqrt{d}] := \left\{a + b\sqrt{d} : a, b, \in \mathbb{Z}\right\}$ *(p. 49)*.

$\mathbb{Z}[\frac{1+\sqrt{d}}{2}]$

The ring $\mathbb{Z}[\frac{1+\sqrt{d}}{2}] := \left\{\frac{a+b\sqrt{d}}{2} : a, b, \in \mathbb{Z}, a \equiv b \bmod 2\right\}$ *(p. 49)*.

$\mathbb{Z}v$

The set $\{kv : k \in \mathbb{Z}\}$ (here $v \in \mathbb{R}^n$).

ζ_m

The primitive m-th root of unity $\zeta_m = e^{\frac{2\pi i}{m}}$.

$\zeta(s)$

The Riemann ζ-function $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ *(p. 44)*.

$\zeta_K(s)$

The Dedekind ζ-function of the number field K, $\zeta_K(s) = \sum_{\mathfrak{a} \lhd \mathcal{O}_K} \frac{1}{N(\mathfrak{a})}^s$ *(p. 52)*.

$\zeta(X,s)$

The ζ function attached to the variety X, $\zeta(X,s) = \exp\left(\sum_{m=1}^{\infty} \frac{N_m}{m} q^{-ms}\right)$ *(p. 79)*.

symbols for the chapter on minimal sets

$[a]_m$

The set of all integers congruent to a modulo m, $[a]_m = \{b \in \mathbb{Z} : b \equiv a \bmod m\}$.

$\langle L \rangle$

The generated set of L, i.e., $\langle L \rangle = \{x \in \mathbb{N} : \exists y \in L \text{ such that } y \lhd x\}$ *(p. 139)*.

$\boxed{n}$

The set of all integers that can be written as a sum of n squares: $\boxed{n} = \{a \in \mathbb{Z} : a = x_1^2 + \cdots + x_n^2 \text{ for some } x_i \in \mathbb{Z}\}$.

$\square \bmod m$

$\square \bmod m = \{n \in \mathbb{N} : n \equiv x^2 \bmod m \text{ for some } x \in \mathbb{N}\}$ *(p. 154)*.

$\#n$

The number of digits of n *(p. 138)*.

$x \lhd y$

x is a subsequence of y (when viewing $x, y \in \mathbb{N}$ in their decimal expansion) *(p. 138)*.

$x * y$

The concatenation of x and y, i.e., string that has first the digits from x and then the digits from y *(p. 138)*.

$M * L$

For $M, L \subset \mathbb{N}$ the set $M * L := \{z \in \mathbb{N} : z = x * y, x \in M, y \in L\}$ (sometimes the asterisk is omitted) *(p. 138)*.

mL

For $m \in \mathbb{N}, L \subset \mathbb{N}$ the set $mL := \{m\}L$ *(p. 138)*.

x^{*k}

The k-fold concatenation of x with itself: $x^{*k} = x^{*(k-1)} * x, x^{*1} = x$ *(p. 138)*.

$\{z\}^*$

The set $\{z^{*k} : k \in \mathbb{N}_0\}$ *(p. 138).*

M^*

For $M \subset \{0,1,\ldots,9\}$ this is the set
$M^* = \{x \in \mathbb{N}_0 : (d \triangleleft x, d \in \{0,\ldots,9\}) \Rightarrow d \in M\}$ *(p. 138).*

M^{*I}

The set $M^{*I} := \{x \in M^* : \#x \in I\}$ for $M \subset \{0,1,\ldots,9\}$ and $I \subset \mathbb{N}$ *(p. 138).*

$\{a * b\}^|$

The set $\{a * b\}^| := \{n \in \mathbb{N} : n = a^{*k} * b^{*l}$ for some $k,l \in \mathbb{N}_0\}$ *(p. 159).*

A_k

The set of natural numbers with k digits in base 10: $A_k := \mathbb{N} \cap [10^{k-1}, 10^k)$ *(p. 158).*

$\delta^n(M)$

The sequence of sets defined via $\delta^0(M) = M, \delta(M) = \delta^1(M) := M \backslash \mathcal{S}(M)$, $\delta^{n+1}(M) = \delta(\delta^n(M))$ *(p. 159).*

$\eta^n(M)$

$\eta^n(M) = |\mathcal{S}(\delta^n(M))|$ *(p. 159).*

$\eta(M)$

$\eta(M) = \eta^1(M)$ *(p. 159).*

$\mu_c(n)$

The digit measure $\mu_c(n) = \frac{10}{9} \cdot 10^{-\#n}$ *(p. 168).*

$\mu_g(n)$

The digit measure $\mu_g(n) = 10 \cdot 10^{-2\#n}$ *(p. 168).*

$\mu_h(n)$

The digit measure $\mu_h(n) = \frac{10}{9} \cdot \frac{1}{\#n} 10^{-\#n}$ *(p. 168).*

$\mu_z(n)$

The digit measure $\mu_z(n) = \frac{20}{3\pi^2} \cdot \frac{1}{(\#n)^2} 10^{-\#n}$ *(p. 168).*

$\psi(n)$

The Dedekind ψ-function, $\psi(n) = \prod_{p|n}(p+1)p^{v_p(n)-1}$ *(p. 148)*.

$\mathcal{S}(M)$

The minimal set of M, i.e., $\mathcal{S}(M) = \{m \in M : \{n \in M : n < m, n \triangleleft m\} = \emptyset\}$ *(p. 139)*.

symbols for the chapter on generators in abelian groups

$\langle a \rangle$

For a group A and $a \in A$, $\langle a \rangle$ is the subgroup generated by a.

$(a)_n^\star$

For $a \in \mathbb{Z}/n\mathbb{Z}$ the set $\{ax : 1 \leq x \leq \mathrm{ord}(a), (x, \mathrm{ord}(a)) = 1\}$ *(p. 176)*.

$(a)_A^\star$

For $a \in A \cong \bigtimes_{i=1}^{k} \mathbb{Z}/(m_i\mathbb{Z})$ the set
$\{(a_i x_i) : 1 \leq x_i \leq \mathrm{ord}(a_i), (x_i, \mathrm{ord}(a_i)) = 1 \text{ for all } i\}$ *(p. 178)*.

$\mathrm{atom}(a)$

The generators of the subgroup $\langle a \rangle$: $\mathrm{atom}(a) = \{a' \in G : \langle a' \rangle = \langle a \rangle\}$ *(p. 176, 177)*.

$N_{A;a,b}(c)$

The number of elements in $S_{A;a,b}(c)$, $N_{A;a,b}(c) = |S_{A;a,b}(c)|$ *(p. 178)*.

$N_{n;a,b}(c)$

The number of elements in $S_{n;a,b}(c)$, $N_{n;a,b}(c) = |S_{n;a,b}(c)|$ *(p. 176)*.

$\mathrm{ord}(a)$

For given m the order of a in the group $(\mathbb{Z}/m\mathbb{Z}, +)$, i.e., $\mathrm{ord}(a) = \frac{a}{\gcd(a,m)}$ *(p. 176)*.

$S_{n;a,b}(c)$

The set $\{(u, v) \in \mathrm{atom}(a) \times \mathrm{atom}(b) : u + v = c\}$ for $a, b, c \in \mathbb{Z}/n\mathbb{Z}$ *(p. 175)*.

$S_{A;a,b}(c)$

The set $\{(u, v) \in \mathrm{atom}(a) \times \mathrm{atom}(b) : u + v = c\}$ for a, b, c in an abelian group A *(p. 178)*.

symbols for the chapter on solutions of linear equations

$\langle \mathbf{x}, \mathbf{y} \rangle$

The inner product $\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n$.

∇f

The gradient of the function f.

$\sum^*_{a \bmod q}$

The summation ranges over all a coprime to q *(p. 185)*.

$\sum_{\mathbf{a} \bmod q}$

In the sum, every entry of $\mathbf{a}$ runs independently modulo q *(p. 185)*.

$\prod_{\pi | a}$

For a in a unique factorization domain: The product over all π dividing a in a fixed prime decomposition *(p. 184)*.

$\mathbf{x} \in \mathfrak{a}$

Each component of $\mathbf{x} \in R^n$ lies in the ideal $\mathfrak{a} \lhd R$.

$\mathbf{x} \equiv \mathbf{y} \bmod a$

$x_i \equiv y_i \bmod a$ for each component of $\mathbf{x}, \mathbf{y} \in R^n$.

$a \sim b$

a and b are associates in the ring R, i.e., $a|b$ and $b|a$.

$\#_y(A, m)$

The number $\left|\{a \in [1, m] : \left|\{\mathbf{x} \in (\mathbb{Z}/a\mathbb{Z})^n : A\mathbf{x} = \mathbf{0} \bmod a\}\right| = y\}\right|$ *(p. 195)*.

$\operatorname{diag}(d_1, \ldots, d_r)$

An $m \times n$ matrix with diagonal elements d_i (with $r = \min\{m, n\}$) and zero off-diagonal elements.

$e_q(x)$

$e_q(x) = e(\frac{x}{q})$ *(p. 19, 185)*.

$S_q(\mathbf{c})$

The sum $\sum^*_{a \bmod q} \sum_{\mathbf{b} \bmod q} e_q(aF(\mathbf{b}) + \langle \mathbf{c}, \mathbf{b} \rangle)$ for $q \in \mathbb{N}$ and $\mathbf{c} \in \mathbb{Z}^n$ *(p. 185)*.

S_m^Δ

The sum $\sum_{I\subset\{1,\dots,r\}}(-1)^{|I|}\left\lfloor\frac{m}{\prod_{i\in I}p_i}\right\rfloor$ for $m\in\mathbb{N}$ and $\Delta=p_1^{v_1}\cdots p_r^{v_r}$ *(p. 196)*.

$S_m^\Delta(y)$

The sum $\sum_{I\subset\{1,\dots,r\}}(-1)^{|I|}\left\lfloor\frac{m}{y\cdot\prod_{i\in I}p_i}\right\rfloor$ for $m\in\mathbb{N}, \Delta=p_1\cdots p_r$ and $y|\Delta, y>1$ *(p. 196)*.

$v_\pi(a)$

The exponent of π in a fixed prime decomposition of $a\in R$.

symbols for the chapter on "Lights Out"

$\xrightarrow{(a,b)}$

Press the button in row a and column b on a "Lights Out" board *(p. 202)*.

$A\otimes B$

The Kronecker product of two matrices A and B *(p. 32)*.

A_n

The absolute value of Λ_n, $A_n=|\Lambda_n|$ *(p. 212)*.

BL_n

The $n^2\times n^2$ matrix corresponding to $BLO(n,k)$ *(p. 205)*.

$BLO(n,k)$

The (bounded) "Lights Out" game on an $n\times n$ board with k colors *(p. 202)*.

$f_n(x)$

The n-th Fibonacci polynomial, recursively defined via
$f_{n+1}(x)=xf_n(x)+f_{n-1}(x),\quad f_1(x)=1, f_2(x)=x$ *(p. 209)*.

J_n

The Toeplitz tridiagonal matrix with 1 on the sub-, super-, and maindiagonal *(p. 205)*.

K_n

The circulant matrix with 1 on the sub-, super-, and maindiagonal and in each corner *(p. 220)*.

Λ_n

The product $\Lambda_n = \prod_{j=1}^{n} \left(1 + 4\cos\left(\frac{j\pi}{n+1}\right)\right)$ *(p. 210)*.

UL_n

The $n^2 \times n^2$ matrix corresponding to $ULO(n,k)$ *(p. 219)*.

$ULO(n,k)$

The unbounded "Lights Out" game on an $n \times n$ board with k colors *(p. 219)*.

V_n

The product $V_n = \prod_{\substack{j,l=1 \\ j \neq l}}^{n} \left(1 + 2\cos\left(\frac{j\pi}{n+1}\right) + 2\cos\left(\frac{l\pi}{n+1}\right)\right)$ *(p. 210)*.

Index

References

[ACGM10] S. D. Adhikari, M. N. Chintamani, Geeta, and B. K. Moriya, *The Cauchy-Davenport theorem: various proofs and some early generalizations*, Math. Student **79** (2010), no. 1 – 4, 109 – 116.

[AF98] M. Anderson and T. Feil, *Turning Lights Out with Linear Algebra*, Math. Mag. **71** (1998), no. 4, 300 – 303.

[AL11] W. Aitken and F. Lemmermeyer, *Counterexamples to the Hasse principle*, Amer. Math. Monthly **118** (2011), no. 7, 610 – 628.

[All93] J.-P. Allouche, *Finite automata and arithmetic*, Séminaire Lotharingien de Combinatoire (1993), 1 – 18.

[AP12] R. C. Alperin and B. L. Peterson, *Integral sets and Cayley graphs of finite groups*, Electron. J. Combin. **19** (2012), no. 1, Paper 44, 12 pp.

[Apo76] T. M. Apostol, *Introduction to analytic number theory*, Springer, 1976.

[Apo90] ———, *Modular functions and Dirichlet series in number theory*, 2nd ed., Springer, 1990.

[AS03] J.-P. Allouche and J. Shallit, *Automatic sequences*, Cambridge University Press, 2003.

[Ash03] R. B. Ash, *A course in algebraic number theory*, 2003, `http://www.math.uiuc.edu/~r-ash/ANT.html` (last checked 13.12.2016).

[Bak90] A. Baker, *Transcendental number theory*, 2nd ed., Cambridge University Press, 1990.

[Ban] J. Bannon, *Collection of equivalent forms of Riemann hypothesis*, MathOverflow, `http://mathoverflow.net/q/39944` (last checked 21.11.2016).

[BCRW08] P. Borwein, S. Choi, B. Rooney, and A. Weirathmueller (eds.), *The Riemann hypothesis*, Springer, 2008.

[BDS16] C. Bright, R. Devillers, and J. Shallit, *Minimal elements for the prime numbers*, Exp. Math. **25** (2016), no. 3, 321 – 331.

[Bes14] A. Beshenov, *Algebraic K-theory of number fields*, Master's thesis, Università degli Studi di Milano / Université de Bordeaux, 2014, `http://algant.eu/documents/theses/beshenov.pdf` (last checked 21.11.2016).

[BHV01] Y. Bilu, G. Hanrot, and P. M. Voutier, *Existence of primitive divisors of Lucas and Lehmer numbers*, J. Reine Angew. Math. **539** (2001), 75 – 122.

[BI09] M. Bašić and A. Ilić, *On the clique number of integral circulant graphs*, Appl. Math. Lett. **22** (2009), no. 9, 1406 – 1411.

[BKS17] I. N. Baoulina, M. Kreh, and J. Steuding, *Deleting digits*, Math. Gaz. **101** (2017), no. 550, 60 – 68.

[BL65] P. T. Bateman and M. E. Low, *Prime numbers in arithmetic progressions with difference* 24, Amer. Math. Monthly **72** (1965), no. 2, 139 – 143.

[Blo04] V. Blomer, *Uniform bounds for Fourier coefficients of theta-series with arithmetic applications*, Acta Arith. **114** (2004), no. 1, 1 – 21.

[BM08] J. A. Bondy and U. S. R. Murty, *Graph theory*, 2nd ed., Springer, 2008.

[Bom06] E. Bombieri, *The Riemann hypothesis*, in: The millennium prize problems (J. Carlson, A. Jaffe, and A. Wiles, eds.), Clay Math. Inst., 2006, pp. 107 – 124.

[Boo03] W. M. Boothby, *An introduction to differentiable manifolds and Riemannian geometry*, 2nd ed., Academic Press, 2003.

[Bos05] S. Bosch, *Algebra*, 6th ed., Springer, 2005.

[BP15] W. D. Brownawell and M. A. Papanikolas, *A rapid introduction to Drinfeld Modules, t-Modules and t-Motives*, 2015, `http://www.math.tamu.edu/~map/BanffSurveyRev2.pdf` (last checked 21.11.2016).

[BR96] R. Barua and S. Ramakrishnan, *σ-game, σ^+-game and two-dimensional additive cellular automata*, Theoret. Comput. Sci. **154** (1996), no. 2, 349 – 366.

[Bra43] A. Brauer, *On the non-existence of odd perfect numbers of form $p^\alpha q_1^2 q_2^2 \cdots q_{t-1}^2 q_t^4$*, Bull. Amer. Math. Soc. **49** (1943), 712 – 718.

[Bro] K. Brown, *Sublime numbers*, `http://www.mathpages.com/home/kmath202/kmath202.htm` (last checked 21.11.2016).

[Brü95] J. Brüdern, *Einführung in die analytische Zahlentheorie*, Springer, 1995.

[Car35] L. Carlitz, *On certain functions connected with polynomials in a Galois field*, Duke Math. J. **1** (1935), no. 2, 137 – 168.

[Cas67] J. W. S. Cassels, *Global fields*, in: Algebraic number theory (J. W. S. Cassels and A. Fröhlich, eds.), Thompson Book Company Inc., 1967, pp. 42 – 84.

[CDdS00] C. Caldeira and J. A. Dias da Silva, *The invariant polynomials degrees of the Kronecker sum of two linear operators and additive theory*, Linear Algebra Appl. **315** (2000), no. 1–3, 125 – 138.

[CF] J. B. Conrey and D. W. Farmer, *Equivalences to the Riemann hypothesis*, `https://web.archive.org/web/20120731034246/http://aimath.org/pl/rhequivalences` (last checked 21.11.2016).

[CF97] J. H. Conway and F. Y. C. Fung, *The sensual (quadratic) form*, The Carus Mathematical Monographs, vol. 26, Mathematical Association of America, 1997.

[CFJR01] T. Chinburg, E. Friedman, K. N. Jones, and A. W. Reid, *The arithmetic hyperbolic 3-manifold of smallest volume*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **30** (2001), no. 1, 1 – 40.

[CFKP97] J. W. Cannon, W. J. Floyd, R. Kenyon, and W. R. Parry, *Hyperbolic geometry*, in: Flavors of geometry (S. Levy, ed.), Cambridge University Press, 1997, pp. 59 – 115.

[CG] W. Y. C. Chen and N. S. S. Gu, *Loop deletion for the lamp lighting problem*, `http://www.billchen.org/preprint/lamp/lamp.pdf` (last checked 21.11.2016).

[Cha03] R. Chapman, *Evaluating $\zeta(2)$*, 2003, https://www.uam.es/personal_pdi/ciencias/cillerue/Curso/zeta2.pdf (last checked 21.11.2016).

[Chr79] G. Christol, *Ensembles presque périodiques k-reconnaissables*, Theoret. Comput. Sci. **9** (1979), no. 1, 141 – 145.

[CJ76] J. H. Conway and A. J. Jones, *Trigonometric Diophantine equations (On vanishing sums of roots of unity)*, Acta Arith. **30** (1976), no. 3, 229 – 240.

[CJW06] J. Carlson, A. Jaffe, and A. Wiles (eds.), *The millennium prize problems*, Clay Math. Inst., 2006.

[Cla] P. L. Clark, *Geometry of numbers with applications to number theory*, http://math.uga.edu/~pete/geometryofnumbers.pdf (last checked 21.11.2016).

[Cla10] ———, *Absolute values III: The fundamental in/equality, Hensel and Krasner*, 2010, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.419.5257&rep=rep1&type=pdf (last checked 21.11.2016).

[CLM71] P. Camion, L. S. Levy, and H. B. Mann, *Linear equations over a commutative ring*, J. Algebra **18** (1971), 432 – 446.

[CN79] J. H. Conway and S. P. Norton, *Monstrous moonshine*, Bull. London Math. Soc. **11** (1979), no. 3, 308 – 339.

[Coh64] J. H. E. Cohn, *Square Fibonacci numbers, etc.*, Fibonacci Quart. **2** (1964), 109 – 113.

[Coh06] H. Cohn, *A short proof of the simple continued fraction expansion of e*, Amer. Math. Monthly **113** (2006), no. 1, 57 – 62.

[Coh07a] H. Cohen, *Number theory Vol. I. Tools and Diophantine equations*, Springer, 2007.

[Coh07b] ———, *Number theory Vol. II. Analytic and modern tools*, Springer, 2007.

[Cona] K. Conrad, *Euclidean proofs of Dirichlet's theorem*, http://www.math.uconn.edu/~kconrad/blurbs/gradnumthy/dirichleteuclid.pdf (last checked 21.11.2016).

[Conb] ———, *Examples of Mordell's equation*, http://www.math.uconn.edu/~kconrad/blurbs/gradnumthy/mordelleqn1.pdf (last checked 21.11.2016).

[Conc] ———, *Hensel's lemma*, http://www.math.uconn.edu/~kconrad/blurbs/gradnumthy/hensel.pdf (last checked 21.11.2016).

[Cond] ———, *Ostrowski for number fields*, http://citeseerx.ist.psu.edu/viewdoc/download?rep=rep1&type=pdf&doi=10.1.1.210.7491 (last checked 21.11.2016).

[Cone] ———, *Pfister's theorem on sums of squares*, http://www.math.uconn.edu/~kconrad/blurbs/linmultialg/pfister.pdf (last checked 21.11.2016).

[Conf] ———, *Pythagorean triples*, http://www.math.uconn.edu/~kconrad/blurbs/ugradnumthy/pythagtriple.pdf (last checked 21.11.2016).

[Cong] ———, *Selmer's example*, http://www.math.uconn.edu/~kconrad/blurbs/gradnumthy/selmerexample.pdf (last checked 21.11.2016).

[Conh] ———, *The congruent number problem*, http://www.math.uconn.edu/~kconrad/blurbs/ugradnumthy/congnumber.pdf (last checked 21.11.2016).

[Coni] ———, *The Hurwitz theorem on sums of squares*, http://www.math.uconn.edu/~kconrad/blurbs/linmultialg/hurwitzlinear.pdf (last checked 21.11.2016).

[Conj] ———, *Why are topological ideas so important in arithmetic?*, MathOverflow, http://mathoverflow.net/q/26083 (last checked 21.11.2016).

[Cor] D. Corwin, *What is the difference between a zeta function and an L-function?*, MathOverflow, http://mathoverflow.net/q/6889 (last checked 21.11.2016).

[Cox13] D. A. Cox, *Primes of the form $x^2 + ny^2$. Fermat, class field theory, and complex multiplication.*, 2nd ed., John Wiley & Sons, Inc., 2013.

[CS63] N. Chomsky and M. P. Schützenberger, *The algebraic theory of context-free languages*, in: Computer programming and formal systems (P. Braffort and D. Hirschberg, eds.), Elsevier, 1963, pp. 118 – 161.

[CS04] J. Carlson and D. Stolarski, *The correct solution to Berlekamp's switching game*, Discrete Math. **287** (2004), no. 1 – 3, 145 – 150.

[Dai90] H. Dai, *On the symmetric solutions of linear matrix equations*, Linear Algebra Appl. **131** (1990), 1 – 7.

[DdSH90] J. A. Dias da Silva and Y. O. Hamidoune, *A note on the minimal polynomial of the Kronecker sum of two linear operators*, Linear Algebra Appl. **141** (1990), 283 – 287.

[Del74] P. Deligne, *La conjecture de Weil. I*, Inst. Hautes Études Sci. Publ. Math. **43** (1974), 273 – 307.

[Del02] É. Delabaere, *Ramanujan's summation*, in: Algorithms Seminar, 2001 – 2002 (F. Chyzak, ed.), Institut National de Recherche en Informatique et en Automatique, 2002.

[Deu53] M. Deuring, *Die Zetafunktion einer algebraischen Kurve vom Geschlechte Eins*, Nachr. Akad. Wiss. Göttingen. Math.-Phys. Kl. (1953), 85 – 94, zweite bis vierte Mitteilung 1955 (13 – 42), 1956 (37 – 76), 1957 (55 – 80).

[DG95] I. J. Dejter and R. E. Giudici, *On unitary Cayley graphs*, J. Combin. Math. Combin. Comput. **18** (1995), 121 – 124.

[DGO15] J. F. R. Duncan, M. J. Griffin, and K. Ono, *Moonshine*, Res. Math. Sci. **2** (2015), Art. 11, 57 pp.

[DH91] G. Damamme and Y. Hellegouarch, *Transcendence of the values of the Carlitz zeta function by Wade's method*, J. Number Theory **39** (1991), no. 3, 257 – 278.

[DHPB15] S. Das, K. Halder, S. Pratihar, and P. Bhowmick, *Properties of Farey sequence and their applications to digital image processing*, arXiv:1509.07757 (2015).

[Dia04] Dias da Silva, J. A., *Linear algebra and additive theory*, in: Unusual applications of number theory (M. B. Nathanson, ed.), DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 64, Amer. Math. Soc., 2004.

[Don87] F. J. H. Don, *On the symmetric solutions of a linear matrix equation*, Linear Algebra Appl. **93** (1987), 1 – 7.

[DS05] F. Diamond and J. Shurman, *A first course in modular forms*, Springer, 2005.

[DW01] Y. Dodis and P. Winkler, *Universal configurations in light-flipping games*, Proceedings of the 12th annual ACM/SIAM Symposium on Discrete Algorithms (2001), 926 – 927.

[Dwo60] B. Dwork, *On the rationality of the zeta function of an algebraic variety*, Amer. J. Math. **82** (1960), no. 3, 631 – 648.

[Edw77] H. M. Edwards, *Fermat's last theorem. A genetic introduction to algebraic number theory*, Springer, 1977.

[EES01] H. Eriksson, K. Eriksson, and J. Sjöstrand, *Note on the lamp lighting problem*, Adv. in Appl. Math. **27** (2001), no. 2 – 3, 357 – 366.

[EHH+92] H. D. Ebbinghaus, H. Hermes, F. Hirzebruch, M. Koecher, K. Mainzer, J. Neukirch, A. Prestel, and R. Remmert, *Zahlen*, 3rd ed., Springer, 1992.

[EK40] P. Erdős and M. Kac, *The Gaussian law of errors in the theory of additive number theoretic functions*, Amer. J. Math. **62** (1940), 738 – 742.

[EK98] S. Eliahou and M. Kervaire, *Sumsets in vector spaces over finite fields*, J. Number Theory **71** (1998), no. 1, 12 – 39.

[EK05] _____, *Old and new formulas for the Hopf–Stiefel and related functions*, Expo. Math. **23** (2005), no. 2, 127 – 145.

[Els09] C. Elsner, *Die Hardy-Littlewood-Methode (Kreismethode)*, Vorlesungsskript, 2009, (not publicly available).

[Els10] C. Elsholtz, *A combinatorial approach to sums of two squares and related problems*, in: Additive number theory (D. Chudnovsky and G. Chudnovsky, eds.), Springer, 2010, pp. 115 – 140.

[ES97] P. Erdős and G. N. Sarkozy, *On cycles in the coprime graph of integers*, Electron. J. Combin. **4** (1997), no. 2, Research Paper 8, 11 p.

[ES16] T. Edgar and J. K. Sklar, *A confused electrician uses Smith normal form*, Math. Mag. **89** (2016), no. 1, 3 – 13.

[EW39] P. Erdős and A. Wintner, *Additive arithmetical functions and statistical independence*, Amer. J. Math. **61** (1939), 713 – 721.

[EW11] M. Einsiedler and T. Ward, *Ergodic theory with a view towards number theory*, Springer, 2011.

[Fal83] G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Invent. Math. **73** (1983), no. 3, 349 – 366.

[Fal95] ———, *The proof of Fermat's last theorem by R. Taylor and A. Wiles*, Notices Amer. Math. Soc. **42** (1995), no. 7, 743 – 746.

[FB06] E. Freitag and R. Busam, *Funktionentheorie 1*, 4th ed., Springer, 2006.

[Fir10] A. Firicel, *Subword complexity and Laurent series with coefficients in a finite field*, arXiv:1001.2548 (2010).

[Fis94] G. Fischer, *Ebene algebraische Kurven*, Vieweg, 1994.

[FL05] M. Fiore and T. Leinster, *Objects of categories as complex numbers*, Adv. Math. **190** (2005), no. 2, 264 – 277.

[Fog02] N. P. Fogg, *Substitutions in dynamics, arithmetics and combinatorics*, Springer, 2002.

[For08] O. Forster, *Der Drei-Quadrate-Satz von Gauß*, in: Vorlesungsskript zu Mathematische Miszellen, 2008, `http://www.mathematik.uni-muenchen.de/~forster/v/misc/threesq.pdf` (last checked 21.11.2016).

[Fra24] J. Franel, *Les suites de Farey et le problème des nombres premiers*, Nachr. Ges. Wiss. Göttingen. Math.-Phys. Kl. (1924), 198 – 201.

[Fri07] E. M. Friedlander, *An introduction to K-theory*, Lectures given at the school on algebraic *K*-theory and its applications, 2007, `http://users.ictp.it/~pub_off/lectures/lns023/Friedlander/Friedlander.pdf` (last checked 21.11.2016).

[FS89] P. Fishburn and N. Sloane, *The solution to Berlekamp's switching game*, Discrete Math. **74** (1989), no. 3, 263 – 290.

[Gan06a] T. Gannon, *Monstrous moonshine: the first twenty-five years*, Bull. London Math. Soc. **38** (2006), no. 1, 1 – 33.

[Gan06b] ———, *Moonshine beyond the Monster*, Cambridge Monographs on Mathematical Physics, 2006.

[GBGL08] T. Gowers, J. Barrow-Green, and I. Leader (eds.), *The Princeton companion to mathematics*, Princeton University Press, 2008.

[Gha99] E. Ghate, *Vandiver's conjecture via K-theory*, Summer school on cyclotomic fields, 1999, `http://www.math.tifr.res.in/~eghate/vandiver.pdf` (last checked 21.11.2016).

[GHK07] H. Gruber, M. Holzer, and M. Kutrib, *The size of Higman-Haines sets*, Theoret. Comput. Sci. **387** (2007), no. 2, 167 – 176.

[GHK09] _____, *More on the size of Higman-Haines sets: effective constructions*, Fund. Inform. **91** (2009), no. 1, 105 – 121.

[GK91] S. W. Graham and G. Kolesnik, *van der Corput's method of exponential sums*, Cambridge University Press, 1991.

[GK97] J. Goldwasser and W. Klostermeyer, *Maximization versions of "lights out" games in grids and graphs*, Congr. Numer. **126** (1997), 99 – 111, `https://www.unf.edu/~wkloster/fibonacci/congnum.ps` (last checked 21.11.2016).

[GKT97] J. Goldwasser, W. Klostermeyer, and G. Trapp, *Characterizing switch-setting problems*, Linear and Multilinear Algebra **43** (1997), no. 1 – 3, 121 – 135.

[Gos96] D. Goss, *Basic structures of function field arithmetic*, Springer, 1996.

[GPZ98] J. Gebel, A. Pethö, and H. G. Zimmer, *On Mordell's equation*, Compositio Math. **110** (1998), no. 3, 335 – 367.

[GR01] C. Godsil and G. Royle, *Algebraic graph theory*, Springer, 2001.

[GR15] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 8th ed., Elsevier/Academic Press, 2015.

[Gra] R. Gray, *Toeplitz and cirulant matrices: A review*, `ee.stanford.edu/~gray/toeplitz.pdf` (last checked 21.11.2016).

[Gri82] R. L. Griess, *The friendly giant*, Invent. Math. **69** (1982), no. 1, 1 – 102.

[Gro65] A. Grothendieck, *Formule de Lefschetz et rationalité des fonctions L*, Séminaire Bourbaki **9** (1965), 41 – 55.

[Hal97] K. Halupczok, *Euklidische Zahlkörper*, Diplomarbeit, Universität Konstanz, 1997, `http://wwwmath.uni-muenster.de/u/karin.halupczok/Publik/diplo.pdf` (last checked 21.11.2016).

[Hat] A. Hatcher, *Topology of numbers*, `https://www.math.cornell.edu/~hatcher/TN/TNbook.pdf` (last checked 21.11.2016).

[HB84] D. R. Heath-Brown, *Fermat's two squares theorem*, Invariant **11** (1984), 3 – 5.

[HB96] ———, *A new form of the circle method, and its application to quadratic forms*, J. Reine Angew. Math. **481** (1996), 149 – 206.

[Hel12] H. A. Helfgott, *Minor arcs for Goldbach's problem*, arXiv:1205.5252 (2012).

[Hel13] ———, *Major arcs for Goldbach's problem*, arXiv:1305.2897 (2013).

[HH12] I. Hilgert and J. Hilgert, *Mathematik - Ein Reiseführer*, SpringerSpektrum, 2012.

[Hig52] G. Higman, *Ordering by divisibility in abstract algebras*, Proc. London Math. Soc. **2** (1952), 326 – 336.

[Hil09] D. Hilbert, *Beweis für die Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl n-ter Potenzen (Waringsches Problem)*, Math. Ann. **67** (1909), no. 3, 281 – 300.

[Hit] N. Hitchin, *Geometry of surfaces*, `https://people.maths.ox.ac.uk/hitchin/hitchinnotes/hitchinnotes.html` (last checked 21.11.2016).

[HJ08] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*, 10th ed., Cambridge University Press, 2008.

[HJM15] Y.-H. He, V. Jejjala, and D. Minic, *From Veneziano to Riemann: A string theory statement of the Riemann hypothesis*, arXiv:1501.01975 (2015).

[HM72] P. Hagis and W. L. McDaniel, *A new result concerning the structure of odd perfect numbers*, Proc. Amer. Math. Soc. **32** (1972), 13 – 15.

[HMU02] J. E. Hopcroft, R. Motwani, and J. D. Ullman, *Einführung in die Automatentheorie, Formale Sprachen und Komplexitätstheorie*, 2nd ed., Addison-Wiley, 2002.

[HSSM05] R. A. Horn, V. V. Sergeichuk, and N. Shaked-Monderer, *Solution of linear matrix equations in a *congruence class*, Electron. J. Linear Algebra **13** (2005), 153 – 156.

[Hur98] A. Hurwitz, *Ueber die Composition der quadratischen Formen von belibig vielen Variablen*, Nachr. Ges. Wiss. Göttingen. Math.-Phys. Kl. (1898), 309 – 316.

[Hus04] D. Husemöller, *Elliptic curves*, 2nd ed., Springer, 2004.

[HW08] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 6th ed., Oxford University Press, 2008.

[IB10] A. Ilić and M. Bašić, *On the chromatic number of integral circulant graphs*, Comput. Math. Appl. **60** (2010), no. 1, 144 – 150.

[Ili09] A. Ilić, *The energy of unitary Cayley graphs*, Linear Algebra Appl. **431** (2009), no. 10, 1881 – 1889.

[IR92] K. Ireland and M. Rosen, *A classical introduction to modern number theory*, 2nd ed., Springer, 1992.

[Iwa97] H. Iwaniec, *Topics in classical automorphic forms*, Amer. Math. Soc., 1997.

[Jag] W. Jagy, *Primes of the form $x^2 + ny^2$*, Mathematics Stack Exchange, `http://math.stackexchange.com/q/954249` (last checked 21.11.2016).

[Jam] James, *Determinant of block tridiagonal matrices*, MathOverflow, `http://mathoverflow.net/q/210492` (last checked 21.11.2016).

[Jos08] J. Jost, *Riemannian geometry and geometric analysis*, 5th ed., Springer, 2008.

[Kan] E. Kani, *Idoneal numbers and some generalizations*, `http://www.mast.queensu.ca/~kani/papers/idoneal-f.pdf` (last checked 21.11.2016).

[Kat07] S. Katok, *p-adic analysis compared with real*, Amer. Math. Soc., 2007.

[Kat10] ———, *Fuchsian groups, geodesic flows on surfaces of constant negative curvature and symbolic coding of geodesics*, in: Homogeneous flows, moduli spaces and arithmetic, Clay Math. Proc., vol. 10, Amer. Math. Soc., 2010.

[KK07] M. Koecher and A. Krieg, *Elliptische Funktionen und Modulformen*, 2nd ed., Springer, 2007.

[Kle08] A. Klenke, *Wahrscheinlichkeitstheorie*, 2nd ed., Springer, 2008.

[Klo] W. Klostermeyer, *Lights Out!: A survey of parity domination in grid graphs*, `https://www.unf.edu/~wkloster/termpaper.pdf` (last checked 21.11.2016).

[KM76] C. G. Khatri and S. K. Mitra, *Hermitian and nonnegative definite solutions of linear matrix equations*, SIAM J. Appl. Math. **31** (1976), no. 4, 579 – 585.

[Kno22] K. Knopp, *Theorie und Anwendung der Unendlichen Reihen*, Springer, 1922.

[Kob93] N. Koblitz, *Introduction to elliptic curves and modular forms*, 2nd ed., Springer, 1993.

[KP01] B. Kalantari and T. H. Pate, *A determinantal lower bound*, Linear Algebra Appl. **326** (2001), no. 1 – 3, 151 – 159.

[Kre15a] M. Kreh, *Adding generators in abelian groups*, Journal for Algebra and Number Theory Academia **5** (2015), no. 3, 81 – 98.

[Kre15b] ———, *Minimal sets*, J. Integer Seq. **18** (2015), no. 5, Article 15.5.3, 38 pp.

[Kre16] ———, *On the number of solutions of linear equations over factor rings of PID*, JP Journal of Algebra, Number Theory and Applications **38** (2016), no. 3, 295 – 318.

[Kre17] ———, *"Lights Out" and variants*, Amer. Math. Monthly (2017), to appear.

[KS07] W. Klotz and T. Sander, *Some properties of unitary Cayley graphs*, Electron. J. Combin. **14** (2007), no. 1, Research Paper 45, 12 pp.

[KS12] ———, *Distance powers and distance matrices of integral Cayley graphs over abelian groups*, Electron. J. Combin. **19** (2012), no. 4, Paper 25, 8 pp.

[KSW99] I. Kovacs, D. S. Silver, and S. G. Williams, *Determinants of commuting-block matrices*, Amer. Math. Monthly **106** (1999), no. 10, 950 – 952.

[Küh49] U. Kühnel, *Verschärfung der notwendigen Bedingungen für die Existenz von ungeraden vollkommenen Zahlen*, Math. Z. **52** (1949), 202 – 211.

[Kuk] A. Kuku, *Introduction to K-theory and some applications*, https://www.math.ksu.edu/~zlin/kuku/Intro-Kthy.pdf (last checked 21.11.2016).

[KX10] M. P. Knapp and C. Xenophontos, *Numerical analysis meets number theory: using rootfinding methods to calculate inverses mod p^n*, Appl. Anal. Discrete Math. **4** (2010), no. 1, 23 – 31.

[KY96] S. Kanemitsu and M. Yoshimoto, *Farey series and the Riemann hypothesis*, Acta Arith. **75** (1996), no. 4, 351 – 374.

[Lan24] E. Landau, *Bemerkungen zu der obenstehenden Abhandlung von J. Franel*, Nachr. Ges. Wiss. Göttingen. Math.-Phys. Kl. (1924), 202 – 206.

[Lau08] M. Laurent, *Linear forms in two logarithms and interpolation determinants II*, Acta Arith. **133** (2008), no. 4, 325 – 348.

[Lee97] J. M. Lee, *Riemannian manifolds*, Springer, 1997.

[Leu96] A. Leutbecher, *Zahlentheorie*, Springer, 1996.

[LMF] The LMFDB Collaboration, *The L-functions and modular forms database*, http://www.lmfdb.org/L/ (last checked 21.11.2016).

[Loo] A. Loos, *Map of maths*, https://dmv.mathematik.de/index.php/forum/bilder-aus-der-mathematik-math-images/437-map-of-maths (last checked 01.02.2017).

[Lot05] M. Lothaire, *Applied combinatorics on words*, Cambridge University Press, 2005.

[LR11] Á. Lozano-Robledo, *Elliptic curves, modular forms, and their L-functions*, Amer. Math. Soc., 2011.

[LS12] C. Li and C. Sia, *Knots and primes*, 2012, http://pub.math.leidenuniv.nl/~lyczakjt/seminar/knots2016-files/knots_and_primes.pdf (last checked 21.11.2016).

[Luc78] É. Lucas, *Théorie des Fonctions Numériques Simplement Périodiques*, Amer. J. Math. **1** (1878), no. 2, 184 – 196.

[Luc90] ———, *Sur les nombres parfaits*, Mathesis **41** (1890), 74 – 76.

[Mah53] K. Mahler, *On the approximation of π*, Indagationes Math. **15** (1953), 30 – 42.

[Mar] M. Marcolli, *Number theory in physics*, http://empslocal.ex.ac.uk/people/staff/mrwatkin/zeta/Marcolli-NTphysics.pdf (last checked 21.11.2016).

[Mar12] W. Marszalek, *Circuits with oscillatory hierarchical Farey sequences and fractal properties*, Circuits Systems Signal Process. **31** (2012), no. 4, 1279 – 1296.

[Mil97] J. S. Milne, *Modular functions and modular forms*, 1997, http://www.jmilne.org/math/CourseNotes/MF110.pdf (last checked 21.11.2016).

[Mil13] ———, *Lectures on étale cohomology*, 2013, http://www.jmilne.org/math/CourseNotes/LEC.pdf (last checked 21.11.2016).

[Möl] M. Möller, *Ergodentheorie*, Skript zur Vorlesung, https://www.uni-frankfurt.de/52367844/ergodentheorie_skript.pdf (last checked 21.11.2016).

[Mol08] L. G. Molinari, *Determinants of block tridiagonal matrices*, Linear Algebra Appl. **429** (2008), no. 8 – 9, 2221 – 2226.

[Mor66] L. J. Mordell, *The infinity of rational solutions of $y^2 = x^3 + k$*, J. London Math. Soc. **41** (1966), 523 – 525.

[Mor05] P. J. Morandi, *The Smith normal form of a matrix*, 2005, http://sierra.nmsu.edu/morandi/notes/SmithNormalForm.pdf (last checked 21.11.2016).

[MPS] Great Internet Mersenne Prime Search, *List of known Mersenne prime numbers*, http://www.mersenne.org/primes/ (last checked 21.11.2016).

[MR03] C. Maclachlan and A. W. Reid, *The arithmetic of hyperbolic 3-manifolds*, Springer, 2003.

[MSC] Mathematics Subject Classification, *The MSC2010 database*, http://www.ams.org/msc/pdfs/classifications2010.pdf (last checked 21.11.2016).

[MT06] M. R. Murty and N. Thain, *Prime numbers in certain arithmetic progressions*, Funct. Approx. Comment. Math. **35** (2006), 249 – 259.

[Nat96a] M. B. Nathanson, *Additive number theory. Inverse problems and the geometry of sumsets*, Springer, 1996.

[Nat96b] ———, *Additive number theory. The classical bases*, Springer, 1996.

[Nat04] M. B. Nathanson (ed.), *Unusual applications of number theory*, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 64, Amer. Math. Soc., 2004.

[Neu92] J. Neukirch, *Algebraische Zahlentheorie*, Springer, 1992.

[New78] M. Newman, *Combinatorial matrices with small determinants*, Canadian J. Math. **30** (1978), no. 4, 756 – 762.

[Niv44] I. Niven, *An unsolved case of the Waring problem*, Amer. J. Math. **66** (1944), 137 – 143.

[NPR13] S. Noschese, L. Pasquini, and L. Reichel, *Tridiagonal Toeplitz matrices: properties and novel applications*, Numer. Linear Algebra Appl. **20** (2013), no. 2, 302 – 326.

[NTW] Tools on Number Theory Web, *Data on Mordell's curve for* $|k| \leq 10\,000$, `http://tnt.math.se.tmu.ac.jp/simath/MORDELL/` (last checked 21.11.2016).

[OEI] OEIS Foundation Inc., *The on-line encyclopedia of integer sequences, sequence A162698*, `https://oeis.org/A162698`.

[O'M73] O. T. O'Meara, *Introduction to quadratic forms*, 3rd ed., Springer, 1973.

[Oss] B. Osserman, *A concise account of the Weil conjectures and étale cohomology*, `https://www.math.ucdavis.edu/~osserman/classes/256B/notes/sem-weil.ps` (last checked 21.11.2016).

[Pav] Algebraic Pavel, *Determinant of block tridiagonal matrices*, Mathematics Stack Exchange, `http://math.stackexchange.com/q/1307671` (last checked 21.11.2016).

[Pfi65] A. Pfister, *Zur Darstellung von* -1 *als Summe von Quadraten in einem Körper*, J. London Math. Soc. **40** (1965), 159 – 165.

[Pil] F. Pillichshammer, *Zahlentheoretische Methoden in der Numerik*, Skript zur Vorlesung, `http://web.maths.unsw.edu.au/~josefdick/preprints/Pillichshammer_lecture_notes.pdf` (last checked 21.11.2016).

[Pla03] A. Plagne, *Additive number theory sheds extra light on the Hopf-Stiefel $\circ$ function*, Enseign. Math. **49** (2003), no. 1 – 2, 109 – 116.

[Pol74] J. M. Pollard, *A generalisation of the theorem of Cauchy and Davenport*, J. London Math. Soc. **2** (1974), no. 8, 460 – 462.

[PRV00] B. Poonen and F. Rodriguez-Villegas, *Lattice polygons and the number 12*, Amer. Math. Monthly **107** (2000), no. 3, 238 – 250.

[QSS01] A. Quarteroni, R. Sacco, and F. Saleri, *Numerische Mathematik 1*, Springer, 2001.

[qua] quanta, *List of local to global principles*, Mathematics Stack Exchange, `http://math.stackexchange.com/q/34053` (last checked 21.1.2016).

[Rad43] H. Rademacher, *On the expansion of the partition function in a series*, Ann. of Math. **44** (1943), no. 2, 416 – 422.

[Rib90] K. A. Ribet, *From the Taniyama-Shimura conjecture to Fermat's last theorem*, Ann. Fac. Sci. Toulouse Math. **11** (1990), no. 1, 116 – 139.

[Rib99] P. Ribenboim, *Fermats last theorem for amateurs*, Springer, 1999.

[Rob99] A. M. Robert, *A course in p-adic analysis*, Springer, 1999.

[RR94] A. Ramsay and R. D. Richtmyer, *Introduction to hyperbolic geometry*, Springer, 1994.

[RV99] D. Ramakrishnan and R. J. Valenza, *Fourier analysis on number fields*, Springer, 1999.

[Sal08] V. Kh. Salikhov, *On the irrationality measure of π*, Russian Math. Surveys **63** (2008), no. 3, 570 – 572.

[Scha] A. Scheffler, *How to solve Lights Out puzzles*, `http://www.hamusutaa.com/pilot/solution.html` (last checked 21.11.2016).

[Schb] J. Scherphuis, *The mathematics of Lights Out*, `http://www.jaapsch.net/puzzles/lomath.htm` (last checked 21.11.2016).

[Sch12] I. Schur, *Über die Existenz unendlich vieler Primzahlen in einigen speziellen arithmetischen Progressionen*, Sitzungsberichte der Berliner Mathematischen Gesellschaft **11** (1912), 40 – 50.

[Sch00] S. H. Schanuel, *Objective number theory and the retract chain condition*, J. Pure Appl. Algebra **154** (2000), no. 1 – 3, 295 – 298.

[Sch03] R. Schoof, *Class numbers of real cyclotomic fields of prime conductor*, Math. Comp. **72** (2003), no. 242, 913 – 937.

[Sch07a] A. Schmidt, *Einführung in die algebraische Zahlentheorie*, Springer, 2007.

[Sch07b] R. Schoof, *Catalans' conjecture*, Springer, 2007.

[Sel51] E. S. Selmer, *The diophantine equation* $ax^3 + by^3 + cz^3 = 0$, Acta Math. **85** (1951), 203 – 362.

[Ser73] J.-P. Serre, *A Course in arithmetic*, Springer, 1973.

[Ser13] C. Series, *Hyperbolic geometry*, 2013, `https://homepages.warwick.ac.uk/~masbb/Papers/MA448.pdf` (last checked 21.11.2016).

[Ser15] ———, *Continued fractions and hyperbolic geometry*, Loughborough LMS Summer School, 2015, `http://homepages.warwick.ac.uk/~masbb/HypGeomandCntdFractions-2.pdf` (last checked 21.11.2016).

[SF07] H. Scheid and A. Frommer, *Zahlentheorie*, 4th ed., Spektrum Akademischer Verlag, 2007.

[SH11] D. Schumayer and D. A. W. Hutchinson, *Physics of the Riemann hypothesis*, Reviews of Modern Physics **83** (2011), 307 – 330.

[Sha00a] J. O. Shallit, *Minimal primes*, J. Recreat. Math. **30** (2000), 113 – 117.

[Sha00b] D. B. Shapiro, *Compositions of quadratic forms*, de Gruyter, 2000.

[Sha04] J. O. Shallit, *Formal languages and number theory*, in: Unusual applications of number theory (M. B. Nathanson, ed.), DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 64, Amer. Math. Soc., 2004.

[Sil99] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, 2nd ed., Springer, 1999.

[Sil08] ———, *The arithmetic of elliptic curves*, 2nd ed., Springer, 2008.

[Sin] D. Singmaster, *Sources in recreational mathematics - an annotated bibliography*, http://puzzlemuseum.com/singma/singma6/SOURCES/singma-sources-edn8-2004-03-19.htm (last checked 21.11.2016).

[SK04] L. Somer and M. Křížek, *On a connection of number theory with graph theory*, Czechoslovak Math. J. **54** (2004), no. 2, 465 – 485.

[So06] W. So, *Integral circulant graphs*, Discrete Math. **306** (2006), no. 1, 153 – 158.

[Sol] *Solving Lights Out*, http://web.archive.org/web/20100704161251/http://www.haar.clara.co.uk/Lights/solving.html (last checked 21.11.2016).

[Sou08] C. Soulé, *Higher K-theory of algebraic integers and the cohomology of arithmetic groups*, 2008, http://www.ihes.fr/~soule/soulehangzhou.pdf (last checked 21.11.2016).

[SPB12] D. Stevanovic, M. Petkovic, and M. Basic, *On the diameter of integral circulant graphs*, Ars Combin. **106** (2012), 495 – 500.

[Spi82] R. Spigler, *An application of group theory to matrices and to ordinary differential equations*, Linear Algebra Appl. **44** (1982), 143 – 151.

[SS83] J. Sakarovitch and I. Simon, *Subwords*, in: Combinatorics on words (M. Lothaire, ed.), Encyclopedia of mathematics and its applications, Vol. 17, Addison-Wesley, 1983.

[SS09] J. W. Sander and T. Sander, *On the kernel of the coprime graph of integers*, Integers **9** (2009), 569 – 579.

[SS13] ———, *Adding generators in cyclic groups*, J. Number Theory **133** (2013), no. 2, 705 – 718.

[ST94] J. H. Silverman and J. T. Tate, *Rational points on elliptic curves*, 2nd ed., Springer, 1994.

[Sta67] H. M. Stark, *A complete determination of the complex quadratic fields of class-number one*, Michigan Math. J. **14** (1967), 1 – 27.

[Ste] J. Steuding, *Ergodic number theory*, Course Notes, http://www.mathematik.uni-wuerzburg.de/~christ/nihon.pdf (last checked 21.11.2016).

[Ste12] _____, *Sampling the Lindelöf hypothesis with an ergodic transformation*, RIMS Kôkyûroku Bessatsu **B34** (2012), 361 – 381.

[Sut89] K. Sutner, *Linear cellular automata and the Garden-of-Eden*, Math. Intelligencer **11** (1989), no. 2, 49 – 53.

[SY14] C.-F. Sun and Q.-H. Yang, *On the sumset of atoms in cyclic groups*, Int. J. Number Theory **10** (2014), no. 6, 1355 – 1363.

[Tat67] J. T. Tate, *Fourier analysis in number fields, and Hecke's zeta-functions*, in: Algebraic number theory (J. W. S. Cassels and A. Fröhlich, eds.), Thompson Book Company Inc., 1967, pp. 305 – 347.

[Ten07] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, 3rd ed., Amer. Math. Soc., 2007.

[Tha] D. S. Thakur, *Automata methods in transcendence*, `https://web.math.rochester.edu/people/faculty/dthakur2/autbanffFinal.pdf` (last checked 21.11.2016).

[Tia02] Y. Tian, *Common solutions of a pair of matrix equations*, Appl. Math. E-Notes **2** (2002), 147 – 154.

[Tis87] M. Tismenetsky, *Determinant of block-Toeplitz band matrices*, Linear Algebra Appl. **85** (1987), 165 – 184.

[Tit86] E. C. Titchmarsh, *The theory of the Riemann zeta-function*, 2nd ed., Oxford University Press, 1986.

[Tom14] R. Tomás, *From Farey sequences to resonance diagrams*, Physical Review Accelerators and Beams **17** (2014), no. 1, p. 0140011.

[Tro] Troglodyte, *Can the repdigit 77...77 be the sum of two squares?*, MathOverflow, `http://mathoverflow.net/q/164791` (last checked 21.11.2016).

[TW95] R. Taylor and A. Wiles, *Ring-theoretic properties of certain Hecke algebras*, Ann. of Math. **141** (1995), no. 3, 553 – 572.

[Vau97] R. C. Vaughan, *The Hardy-Littlewood method*, 2nd ed., Cambridge University Press, 1997.

[Wad41] L. I. Wade, *Certain quantities transcendental over $GF(p^n, x)$*, Duke Math. J. **8** (1941), 701 – 720.

[Wal16] C. Walkden, *Hyperbolic geometry*, 2016, http://www.maths.manchester.ac.uk/~cwalkden/hyperbolic-geometry/hyperbolic_geometry.pdf (last checked 21.11.2016).

[Was97] L. C. Washington, *Introduction to cyclotomic fields*, 2nd ed., Springer, 1997.

[Wat] M. R. Watkins, *A directory of all known zeta functions*, http://empslocal.ex.ac.uk/people/staff/mrwatkin//zeta/directoryofzetafunctions.htm (last checked 21.11.2016).

[Wei49] A. Weil, *Numbers of solutions of equations in finite fields*, Bull. Amer. Math. Soc. **55** (1949), 497 – 508.

[Wei73] P. J. Weinberger, *Exponents of the class group of complex quadratic fields*, Acta Arith. **22** (1973), 117 – 124.

[Wen] M. Wendt, *Why is 12 the smallest weight for which a cusp forms exists*, MathOverflow, http://mathoverflow.net/q/242415 (last checked 21.11.2016).

[Wil95] A. Wiles, *Modular elliptic curves and Fermat's last theorem*, Ann. of Math. **141** (1995), no. 3, 443 – 551.

[Wil06] ———, *The Birch and Swinnerton-Dyer conjecture*, in: The millennium prize problems (J. Carlson, A. Jaffe, and A. Wiles, eds.), Clay Mathematics Institute, Amer. Math. Soc., 2006.

[Yue05] W.-C. Yueh, *Eigenvalues of several tridiagonal matrices*, Appl. Math. E-Notes **5** (2005), 66 – 74.

[Yuz81] S. Yuzvinsky, *Orthogonal pairings of Euclidean spaces*, Michigan Math. J. **28** (1981), no. 2, 131 – 145.

[Zag90] D. B. Zagier, *A one-sentence proof that every prime $p \equiv 1 \pmod{n}$ is a sum of two squares*, Amer. Math. Monthly **97** (1990), no. 2, 144.

[Zag91] ———, *The Birch-Swinnerton-Dyer conjecture from a naive point of view*, in: Arithmetic algebraic geometry, Progr. Math., no. 89, Birkhäuser, 1991, pp. 377 – 389.

[Zag94] ———, *Values of zeta functions and their applications*, in: First european congress of mathematics, vol. II, Progr. Math., no. 120, Birkhäuser, 1994, pp. 497 – 512.

[ZE66] H. Zassenhaus and W. Eichhorn, *Herleitung von Acht- und Sechzehn-Quadrate-Identitäten mit Hilfe von Eigenschaften der verallgemeinerten Quaternionen und der Cayley-Dicksonschen Zahlen*, Arch. Math. **17** (1966), 492 – 496.

[Zha04] X. Zhang, *Hermitian nonnegative-definite and positive-definite solutions of the matrix equation $AXB = C$*, Appl. Math. E-Notes **4** (2004), 40 – 47.

[Zwi05] D. Zwillinger, *Errata for table of integrals, series, and products, 6th edition*, 2005, `http://www.mathtable.com/errata/gr6_errata.pdf` (last checked 21.11.2016).